CAPITAL UNIVERSITY OF SCIENCE AND
TECHNOLOGY, ISLAMABAD

ST_{index} to Rank Awardees using Neural Network

by

Muhammad Saqib

A thesis submitted in partial fulfillment for the
degree of Master of Science

in the

Faculty of Computing
Department of Computer Science

2020

My dissertation work is devoted to My Family, My Teachers and My Friends. I have a special feeling of gratitude for My beloved parents and brothers. Special thanks to my supervisor whose uncountable confidence enabled me to reach this milestone. I would also like to dedicate my work to my mother who has always pushed me to finish my thesis on time

CERTIFICATE OF APPROVAL

ST_{index} to Rank Awardees using Neural Network

by

Muhammad Saqib

(MCS183015)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Ayyaz Hussain	QAU, Islamabad
(b)	Internal Examiner	Dr. Abdul Basit Siddiqui	CUST, Islamabad
(c)	Supervisor	Dr. Muhammad Tanvir Afzal	CUST, Islamabad

Dr. Muhammad Tanvir Afzal
Thesis Supervisor
August, 2020

Dr. Nayyer Masood
Head
Dept. of Computer Science
August, 2020

Dr. Muhammad Abdul Qadir
Dean
Faculty of Computing
August, 2020

Author's Declaration

I, **Muhammad Saqib** hereby state that my MS thesis titled "**$\mathbf{ST_{index}}$ to Rank Awardees using Neural Network**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.

(Muhammad Saqib)
Registration No: MCS183015

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**ST_{index} to Rank Awardees using Neural Network**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been dully acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.

(Muhammad Saqib)
Registration No: MCS183015

Acknowledgements

Say, "He is Allah, [who is] One. Allah, the Eternal Refuge. He neither begets nor is born. Nor is there to Him any equivalent" Al-Quran [112:1-4]. I would like to say ***Alhamdulillah***, for blessing me, with strength to finish this work.

My heartfelt thanks to my parents for their love, prayers and everything what I need. A special thanks to my Brothers for their support. Last but not the least my friends, with whom this journey proved to be a garden with full of flowers to me.

Special thanks to my esteemed supervisor ***Dr. Muhammad Tanvir Afzal*** for the support, assistance, inspiration and guidance in the field of research. I have learned not only research under his supervision, but also attitude towards others. Sir "You are one of those persons in my life, whom I love to follow".

May Allah shower his countless blessing upon all of you, ***JazakAllah***

(Muhammad Saqib)

Registration No: MCS183015

Abstract

Researchers are ranked in order to determine the significance of their work in the scientific community. Qualitative judgements employs subjective evaluation based on unquantifiable information. There is not any established way to assess the quality of work and we do not have any criteria for qualitative assessment. More than 50 quantitative parameters have been identified by the scientific community up to the date, including publication count, citation count, h-index and its variants. The current state-of-the-art in authors ranking does not determine the best parameter that effectively maps on experts' qualitative assessment. These parameters are proposed taking clever scenarios into consideration. In these scenarios it is very difficult to determine the significance and effect of each parameter over the other. They must therefore be tested in unconscionable situations.

Such parameters were tested on a domain specific dataset in state-of-the-art literature. Researchers use these criteria without recognizing the infallibility of the parameter over others for evaluating individual researchers, research groups etc. There is not any standard available to rank these parameters. The effectiveness of author evaluation parameters over other parameters could be difficult to evaluate. Therefore, such parameters need to be assessed comparatively on a comprehensive dataset from various fields. The purpose of this study is to propose an index which could better map the prestigious researchers to qualitative judgments of experts. We have considered datasets from three fields, Civil Engineering, Mathematics and Neuroscience. Each of the above three data set contain 250 non-award winners and 250 award winners from prestigious scientific societies of respective field. To make the proposed methodology more extensive, we also have compiled all three datasets to find the more comprehensive rank of each parameter among the primitive, citation-intensity and age-based parameters. The complete dataset, contain 1500 researchers, 376,963 publications and their 68,049,621 citations. Afterward, weights of parameters has been analyzed against the ranked lists of researchers to find top five ranked parameters by the Neural network. Secondly, this research focuses on to identify the correlation of each parameter for each field. We have identified the top two parameters for each domain as well as for comprehensive

dataset.

It has been discovered during the analysis of primitive, citation-intensity and age-based parameters that, the primitive Author/papers and A-index have a strong impact on the award recipients of Civil Engineering domain, and for Mathematics Author/papers and f-index, and for Neuroscience domain it has been observed that K-index and A-index proved its strong association with the award recipients of the respective fields. Author/papers and f-index found to have a strong relation when, we combined all three datasets. Afterward, we have analyzed the trend of awardees and following findings have been observed; In ranked list of top 100 researchers by the Authors/paper and A-index parameter, 85% are awardees. For Mathematics and Neuroscience, in list of top 100 researchers 83% and 85% are award recipients respectively. In combined dataset, 80% occurrences of award recipients found in ranked list of 100 researchers by Author/papers and f-index. We have proposed ST_{index} based on these findings for each domain and on a comprehensive dataset as well.

Contents

List of Figures

List of Tables

Abbreviations

ACI	American Concrete Institute
AMS	American Mathematics Society
ANS	Australian Neuroscience Society
ASCE	American Society of Civil Engineers
CSCE	Canadian Society for Civil Engineering
CNS	Cognitive Neuroscience Society
FENS	Federation of European Neuroscience societies
ICE	Institution of Civil Engineers
IMU	International Mathematical Union
IBRO	International Brain Research Organization
LMS	London Mathematical Society
SFN	Society for Neuroscience

Chapter 1

Introduction

Researchers are ranked in the scientific community to assess the importance of their research. According to James and Raheel [1][2], ranking helps to answer questions such as who deserves a grant or award? Whose work makes an impact? Who will get a Scientific Society Fellowship and Membership? The answer to the above questions is oriented towards the qualitative effect of the researcher. Every year, internationally renowned scientific societies nominate award winners based on their qualitative impact. The quality of one's work is not measured by criteria. One way to assess the consistency of the research is to map the quantitative parameters to qualitative judgments. A variety of quantitative techniques for the research evaluation have been proposed. With the exponential growth of research papers, including the information and citations of millions of authors. Assessing the impact of the researcher on such a vast amount of information has become increasingly challenging [3].

1.1 Background

The conventional way for assessing the impact of the researcher was the number of publications [4]. The number of research articles published were used to be counted in order to determine the impact of the author's work on the scientific

community. It is not necessarily valid to conclude that the number of publications is adequate to assess the effect of the researcher without considering the publication venue. Cameron explained the problem with an example [5]. He selected two researchers in the database field. One of them has a long list of publications, while the other has fewer. One of them is E.F. Codd who is the inventor of the relational database. E.F. Codd considered more prolific than Hector Garcia Molina, as he also won the Turing award twice (1981, 1994), however, E. F. Codd has only 49 publications. 'Hector Garcia-Molina,' on the other hand, has 248 publications by the conclusion of the year when E.F.Codd died. This example shows a situation in which, Hector Molina will be ranked quite higher than E. F. Codd and this might not be acceptable by the scientific community by considering the impact of E. F. Codd on database field. In such a situation, the scientific community can be seriously impacted by the number of publications. Therefore, a new technique i.e. citation count was introduced to fix this issue [6].The number of citations earned by the researcher is known to be the impact of the researcher in the scientific community. The more endorsements the researchers have provided, the more highly profiled they are considered. However, the proposed solution proved to be inefficient, as (i) researchers may illegitimately increase their citations, (ii) researchers may cite to criticize a paper, and (iii) survey papers typically have more citations. All of these factors reduce the chances of innovative work getting higher rank.

In order to assess the true performance of the researcher, it was appropriate to include both publications and citations. Therefore, in 2005, a technique i.e. h-index was introduced by Jorge Hirsh, in which he represented both quantity and the quality of publications by a single number [7]. It is a very critical test to determine the impact of researchers. This approach became successful and opened up a way for scientific community towards a new area of study. Hirsh has made significant changes to the H-index until the masses consider it as a benchmark for evaluating the skills of the scientific community.

The H-index is a quantitative indicator, but the H-index omitted a variety of issues related to publications and their citations. The h-index is used throughout the world and had made computation simple. Its implications have been observed by

many and is now the most common standard adapted by research institutions. By 4 June 2020, the initial h-index article had 9,314 citations. Hirsh claims that the h-index measures precisely whether a researcher has received a prestigious award, such as the President of the National Academy or the Nobel Prize. He measured the h-index for 10 well-known bio scientists and has identified higher h-index for the authors who were cited more [8]. Nevertheless, Dienes noted that the h-index has some drawbacks. One of its dimensions notes that further increase in citations to the index documents does not add to the researchers' impact [9]. There are several circumstances in which the h-index may provide incorrect details about a scientist's impact. For example, consider three researchers' data with the papers and the citations as shown in 1.1. Without taking into account the distinctive performance of X3, the h-index = 2 for all three researchers in the given scenario. Consider another example of three researchers with the following data 1.2. Researcher X1 having h=10 and the other two researchers gained h=2, even though their publications and citation count is same.

This shows the h-index deficiency as it shows that citation and publication are considered by the h-index in two separate dimensions. The h-index did not explain how the conversion value converges between two opposite dimensional values [9]. The h-index also does not take into account the social aspect swhich is considered to be the key factor in identifying potential experts from any domain [10]. The researchers have been directed into the shortcomings of the h-index and they have suggested new indices to fix h-index deficiencies. For example, g-index [11], A-index [12], R-index [13], hg-index [14], Ar-index [13] etc. All these parameters depend entirely on publications, citations and their age and are divided into three major categories. i) Primitive parameters, ii) Citation-intensity-based parameters and iii) age-based parameters [15]. Such techniques are suggested ingeniously. These approaches have also been tested on different datasets. It was difficult to understand the significance and effectiveness of each strategy in comparison to the other. Scientists consider new ways to think about existing methods. The rise in the assessment work of these indices rapidly explored their positives, weaknesses and vulnerabilities [2]. From the list of more than 50 indices proposed by the

TABLE 1.1: The Misleading h-index I

Publications	X1's Citations	X2's Citations	X3's Citations
1	6	10	700
2	4	10	500
3	1	1	1
4	1	1	1
5	1	1	1
6	0	1	1
7	0	1	1
8	0	0	1
9	0	0	1
10	0	0	0

scientific community, one index is not agreed upon yet to assess the ranking of researchers. To determine an approach that produces reliable results when assessing the impact of the researcher, a detailed review of all these measures is required. The award winners are considered to be benchmarks in previous studies to determine the performance of these parameters in scientific society [16][17]. The researchers have been ranked by these indices and the occurrences of winners have been manually been computed into a ranking. On the basis of award recipient count, they have ranked these indices [2][16][17]. However, the most efficient measurement parameter is still unclear. To know the most successful authors' ranking parameter, which provides the best possible outcomes, these parameters should be weighed using the data set of various fields of study through current machine learning techniques.

Therefore, this research aims to use machine learning technique, to find weights of the primitive, Citation-intensity, and age-based parameters by means of their effectiveness for award recipients. Primitive parameters include: i) Total number of publications, ii) Total number of citations, iii) Total number of devoted years in research, iv) Authors/Paper, v) Cites/Year, vi) Cites/Paper, vii) Citations in H-Core, and viii) h-index. Citation-intensity-based parameters include: i) g-index,

TABLE 1.2: The Misleading h-index II

Publications	X1's Citations	X2's Citations	X3's Citations
1	10	90	50
2	10	10	42
3	10	0	1
4	10	0	1
5	10	0	1
6	10	0	1
7	10	0	1
8	10	0	1
9	10	0	1
10	10	0	1

ii) hg-index, iii) A-Index iv) R-Index, v) f-index, vi) p-index, vii) q2-index, viii) k-index, ix) e-index, and x) hm-index. Age-based-parameters include:i) hc-index, ii) hl-index, iii) hl-norm, iv) AW-index, v) hl-annual, vi) M-quotient, and vii) AR-index.

Secondly, correlation matrix is formed for all the indices to identify the unique features for the nomination of award recipients. To comprehensively evaluate the proposed index three diversified fields have been selected in this thesis such as: (i) Mathematics, (ii) Civil Engineering and, (iii) Neuroscience have been selected. The dataset from each of these fields consist of 250 non awardees and 250 awardees of respective field of study. Awards recipients of prestigious scientific societies of each of the above field has been considered for evaluation. For mathematics, considered prestigious scientific societies are, i) IMU, ii) LMS (London Mathematical Society), and iii) AMS (American Mathematical Society). For neuroscience, considered scientific societies include, i) SFN (Society for Neuroscience), ii) FENS (Federation of European Neuroscience societies), iii) CNS (Cognitive Neuroscience Society), and iv) ANS (Australian Neuroscience Society). For civil engineering following societies have been considered for evaluation of the proposed technique, i) ACI (American Concrete Institute), ii) ASCE (American Society for Civil Engineering), iii) CSCE (Canadian Society for Civil Engineering), and iv) ICE (Institution of

Civil Engineers).

For the purpose of this research, a list of award recipients was first compiled for data collection between 2011 and 2019 from the corresponding sites of the societies listed above. Then the primitive, citation-intensity and age-based h-index variants were measured. After the completion of the computation phases, we have used the neural network to identify these parameters based on its significance for award recipients. At last we have formed correlation matrix to find those parameters which bring different awardees in the top ranking and then those features are used to formulate a new index.

1.2 Problem Statement

In literature, when a new evaluation parameter is introduced, it is evaluated either on hypothetical cases or on clever test data set. There is no benchmark available which can define the superiority of each parameter over the other by taking into account the latest techniques of machine learning. Furthermore, it has been observed in literature that current state-of-the-art indices do not map on qualitative judgements of experts. Problem being focused in this thesis is to propose an index that is based on critical analysis of the literature and empirical studies of available dataset that could be able to improve the possibility to bring more prestigious awardees in the top ranking. Additionally, these techniques have been evaluated on domain specific dataset. Therefore, to propose a new index and to determine which parameter effectively maps on the quality evaluation of experts, an in-depth analysis of all such parameters should be carried out.

1.3 Research Questions

Based on problem identified in previous section, this research aims to propose and evaluate a new author assessment parameter by considering the awardees of

prestigious scientific societies and non-awardees from the field of Civil Engineering, Mathematics and Neuroscience. The following research questions have been formulated in this thesis.

1.3.1 Research Question 1

What are those weights of indices which contribute effectively to rank prestigious researches of respective fields?

1.3.2 Research Question 2

Which expression (index) can produce better results to rank the prestigious researchers as compared to existing indices?

1.4 Purpose

The purpose of this study is to propose an index which map effectively on qualitative judgments of experts in the field of Civil engineering, Mathematics and Neuroscience. In order to assess the proposed index, 250 non-award recipients and 250 award recipients of each category are considered. Several previous studies have also included award recipients as a benchmark for evaluating the assessment parameters [2][16][17].

1.5 Scope

The scope of this thesis is to explore quantitative parameters for author ranking in the field of Civil Engineering, Mathematics and Neuroscience. The dataset and parameters mentioned in section 1.1 is used to find weights and correlation of mentioned parameters based on their effectiveness for awardees, in the fields of Civil

Engineering, Mathematics and Neuroscience. Afterwards, we have formulated an index which effectively maps on the qualitative judgments of experts of respective fields.

1.6 Application of Proposed Approach

This work would benefit different groups of people in the following ways:

1.6.1 Board Members of Scientific Society

The thorough findings of our research would make it possible for board members to determine such as, i) who deserves to be provided society membership, ii) who should have the research project assignments, iii) who ought to be awarded?

1.6.2 Individual Researchers

The research will lead individual researchers to be listed in their respective fields as recipients of awards.

1.6.3 Expert Ranking System

The expert rating system will use this finding to identify researchers in a specific field.

1.7 Limitations

We focused on three areas of experiments, i) Civil Engineering, ii) Mathematics, and iii) Neuroscience. The experiment dataset comprises 500 records from each domain mentioned above. Among those 500, 250 are awardees and 250 are non-awardees. However, all researchers in the respective areas are not covered.

Chapter 2

Literature Review

Each year renowned scientific societies invite researchers, on the basis of their research impacts on the scientific community, to recognize researchers' contribution. There are several ways to honor scientists. For example, a renowned researcher can be selected as a journal editor or research paper reviewer [18]. An important question in the international scientific community is the assessment of the impact of research work. Another argument is to give him, the post-doctoral position or the supervision of the funded projects by examining the impact of the research. With the rapid growth of researchers' research publications, including millions of publication and citations. Analyzing the impact of the researcher on such an enormous amount of information has become increasingly challenging. Various techniques have been suggested in the literature to quantify the influence of researchers, research groups, journals or universities in the scientific community. For example, the total number of published articles and their citations, the average number of co-authors in published articles, and hybrid approaches, such as h-index, etc. However, these techniques have been considered vulnerable in the research impact assessment of the scientific community [19]. It has also been shown that it is extremely difficult to distinguish between two researchers with different objective: one that often contributes to the scientific community by publishing hundreds of research papers annually and, on the other hand, that does not contribute regularly, but focuses on publishing a few in a year, with a high level of novelty [20].

In 2012, Smolinsky and Lercher discussed the impact of different researchers in mathematics and subcategories on the basis of their citations [21]. They argued that, the discrepancy between citations is primarily because of published articles and the way they are linked internally. They discussed that certain mathematical subcategories could obtain a better citation only because they are connected to the highly cited areas [21].

In 2005, Jorge Hirsh presented the h-index to fix the limitations of citations and publications. The h-index helps us measure the impact of a single researcher in the scientific community. Although the h-Index has got high attention from the scientific community as, it's simple and straightforward, it's use is still common in the scientific community [7]. However, researchers has identified number of loop-holes in the traditional h-index.

Costas and Bordons also established the shortcomings of the h-index as a matter of career length for researchers neglected by the h-index. For example, a young researcher who just started a career a year or two ago will have less number of publications and citations. It is obvious that it took some time to increase the number of publications and to influence the scientific community, resulting in a low h-index. A parameter that accommodates the best evaluation of the researcher is therefore necessary [22]. In order to resolve these deficiencies of conventional researchers' assessment techniques, researchers have suggested a number further variants of the h-index.

The rest of this chapter has been categorized into following classes:

(i) Primitive parameters

(ii) Citation-intensity-based-parameters

(iii) Age-based-parameters

Primitive parameters include, i)Total no of publications, ii) Total no of citations, iii) Total no of devoted years in research, iv) Authors/Paper, v) Cites/Year, vi) Cites/Paper, vii) Citations in H-Core, and viii) h-index.

Citation-intensity-based parameters such as, i) g-index, ii) hg-index, iii) A-Index

iv) R-Index, v) f-index, vi) p-index, vii) q2-index, viii) k-index, ix) hm-index and x) e-index.

Age-based parameters include, i) hc-index, ii) hI-index, iii) hInorm, iv) Aw-index, v) hIannual, vi) m-quotient, and vii) AR- index.

Finally, the state-of-the-art literature is reviewed that has analyzed and applied these standards in evaluation systems.

2.1 Primitive Parameters

The ongoing work in academic publications contain information and citations from millions of authors. Ranking potential experts can therefore guide decision-making in any research domain. In state-of-the-art literature, only publications and their citations are considered to be basic criteria for the recognition of researchers [12]. However the rapid growth in the numbers, the average number of publications and citations by papers published for scientific research, which define specific criteria for evaluating the value of scientific work may also be recognized as basic parameters. The conventional approach for finding the most prominent researcher is the total number of research articles published by the researcher. According to the publication count as criterion for impact measurement, a researcher who has more publications than other domain researchers is known to be more profiled [4]. The main aspect of the publication count is that it is well-known and clear proof of the scientific know-how of researchers in the scientific community. However, there is a continuing dispute that only publications cannot be used for the assessment of researchers because they give equal importance to all the authors of the publication even to an author who is at the end of the list with very little contribution. Furthermore, publishing venue matters a lot. Some authors are publishing in high impact venues whiles other might be publishing in low quality conferences. How both of the authors stand on the same scale.

To overcome inefficiencies of publication count, a citation count was suggested for the assessment of the impact of researchers [6]. According to the citation count, if researcher's work has received more citations than other researchers' work, the

TABLE 2.1: The h-index

Publications	Citations
1	9
2	8
3	7
4	7
h(5)	6
6	3

researcher is ranked higher as compare to others. In the scientific world, citation count is commonly used to assess the influence of the researcher. However, there are some limitations in citation counts as well. For example, by doing self-citations, researchers may increase their impact. To address this, it is possible to consider an average number of citations to identify the true impact of researchers that also reduce the impact of self-citations. The high-profiled research papers of researchers have been considered in publications, it was observed in their published articles that they usually have a few authors on average.

Therefore we have also considered the following parameters: (1) the average number of researchers, (2) average number of gained citations by the researcher in year, and (3) number of devoted years for research in field. Hirsh suggested the h-index that evaluate the researchers' impact by combining the publications and their citations in single number [7]. The h-index of a researcher is "x" if at least "x" of his papers have received at least "x" citations. Table 2.1 shows the h-index of researcher, who has 6 publications and their citations listed in descending order. The h-index of the author is 5, as 5 of his papers has achieved at least 5 citations each. The inaccuracy of the h-index is that the rise in the number of index publications does not put any impact to the h-index. Therefore, we also considered the number of the h-core citations as an individual primitive evaluation parameter. In order to reduce the shortcoming of h-index, for example, the value of h-index does not decrease with the age of the author. In addition to their quality of study, new researchers with low publication and citation rates are often less likely to be

ranked as experts. A number of other variants of the h-index have been suggested, which fall within the citation- intensity range based on and age-related variants of the h-index.

2.2 Citation-Intensity-Based Parameters

The h-index excludes the publications and their cittations which are not part of the h-core results in information loss.. To address this weakness, the g-index was introduced by Egghe [11]. The g-index of an author is a publication count which has equal or more sum of citation than square of index number of publication. Table 2.2 shows, the h-index, g-index and hg-index having 6 publications and citations listed for researchers. To combine the impact of both h-index and g-index, the hg-index is suggested which is calculated by taking the square root of the product of h-index and g-index [14]. A-Index considers the rise in h-core citations that the h-index neglects, taking the average number of citations in h-core [12]. Table 2.3 shows, the A-index of researcher having 7 publications and their listed citations. A-index overcomes the information loss by h-index. However, it hurts from another challenge illustrated by the example below. There are two researchers X1 and X2. Where X1 has 30 and X2 has 50 publications. One of the paper of X1, gained 10 citations and rest cited exactly once and one of the paper of X2 is cited 10 times and all other have been cited twice. In the above scenario, according to quantitative measure, X2 looks more prolific than the X1. However, A-index assign higher rank to X1 with score 10 than X2 with score 6 [12]. R-index is introduced to overcome the shortcomings of the A-index [13]. R-index also emphasizes on the importance of h-core citations which are ignored by the h-index, by taking the square root of the number of h-core citations. Table 2.4 shows the R-index of a researcher with 7 publications and listed citations.

The e-index represents the citations in h-core, which are ignored by the original h-index [23]. E-index signifies the loss of information of h-core by the h-index. If h-index considers all the citations of h-core, the loss will be zero, so the value of e-index will also be zero as shown in Table 2.5.

TABLE 2.2: The g-index

Publications	Citations	g^2	Σ
1	20	1	20
2	10	4	30
h(3)	5	9	35
4	0	16	35
g(5)	0	25	35
6	0	36	35
hg-index=$\sqrt{3 \times 5}$=3.87			

TABLE 2.3: The A-index

Publications	Citations	H-core Citations
1	9	9
2	8	17
3	7	24
4	7	31
h(5)	6	37
6	3	A-index=$\frac{37}{5}$=7.4
7	2	

Similar to hg-index, the q2-index also depends on m-index in addition to the h-index [24]. This modification may add benefits, but may also be vulnerable. It can be calculated by taking the square root of the product of h-index and m-index. Whereas the m-index is median of citations present in h-core. It gives a detailed evaluation of the researcher as compared to when h-index or m-index considered individually.

Publication and their citations other than h-core are not considered by h-index. To avoid this loss, k-index is proposed. The k-index also considers the publication and citations that are not in h-core [25]. Whereas, the p-index reaches the ideal combination between the total citations and mean citations [26].

TABLE 2.4: The R-index

Publications	Citations	H-core Citations
1	9	9
2	8	17
3	7	24
4	7	31
h(5)	6	37
6	3	R-index=$\sqrt{37}$=6.08
7	2	

TABLE 2.5: The e-index

Publications	Citations	H-core Citations
1	7	7
2	5	12
3	5	17
4	4	21
h(5)	4	25
6	3	e-index=$\sqrt{25-25}$=0
7	2	

2.3 Age-Based Parameters

The h-index of researchers tends to increase even after the researcher has been inactive in the field for a long time. Age-based variants consider the age of publications as well as reducing the impact of inactive authors, which is ignored by the h-index. In 2007, Jin argues that considering the age of publications is a sufficient condition to be able to assess improvements in performance over time [13]. The Ar-Index can increase or decrease over time. The Ar-Index consists of the number of quotes divided by the year of publication. When the number of years increases and the amount of citations received stays the same, the effect of the researcher decreases naturally. In order to keep the effect strong, researchers need to stay involved in the field. The m-quotient also takes account of the time problem [27].

H-index does not take into account the age of the publication. Some researchers may have contributed many important articles generating a strong h-Index, but are now inactive in research. In comparison, researchers who sustain publishing or beginners who have the highest probability of making substantial contributions in the near future are subject to a low h index due to time constraints. A contemporary h-index (hc-index) considers the weighted age of article [28]. hc-index considers the age of article in term of multiple of 4 years. This indicates that the paper published in the calendar year has been assigned the maximum weight. The more the publication is older, in terms of multiple of 4 years, it assigns the less the weight to that article. The hi-index seeks to evaluate researchers of different ages [29]. It considers the h-index of each year for assessing the impact of researchers. A researcher having only year in research can be compared with the researcher with 30 years in research. However, it only compares the researchers considering the outcome of particular year. The hl-annual also deal with the comparison of researchers with different lengths of career. It refers to the annual average change of the individual h-index [30].

Some of the h-index variants sought to extend the h-core citations, or considered h-tail separately, rather than taking all citations into account. The Aw-index finds that the age of each publication varies from that of the Ar-index, which only considers regularly cited publications [31]. The h-norm shall take into account the normalized citations. It normalizes the number of citations for each publication, and then sets the h-Index for the normalized citations. This is a much more sophisticated version. As normalized quotes were considered which earlier were not included in the initial h-index [32].

2.4 Evaluation of H-variants:

Evaluating the impacts of scientists, research groups, conferences, journals, which are always seen as astonishing. The scientometrics research area focuses on the quantitative effects and the impact of science and scientific research. The three

measures for scientometrics are paper-level metrics, author-level metrics and journal-level metrics. In addition, scientists have developed a range of methods for recognizing the influence of authors, papers, journals and research institutions. This thesis focuses on the effects of individual researchers. There are three most widely used approaches to assess the impact of researchers' research work [33]. Different experts gather views and recommend the best researcher in the first approach. It is almost of a random approach that does not have to be trained before it is applied. The approach is valuable as it incorporates the experiences of the different professionals who evaluate their research work and examine the impact of the researcher. Nevertheless, since this approach is manual, it is vulnerable to objectivity. The number of citations collected by their publications is evaluated by researchers in a second type of approach. The citation count technique is most often used to evaluate citation information by using digital repositories. There are, therefore, fewer chances of human error. The third method is to determine the influence of work on the science community using award winners from leading scientific societies [16][17].

The new cutting edge focuses on certain statistical metrics, including the number of publications, the amount of publications published annually, the number of citations earned, and the average number of citations earned. It never differentiates between low-quality researchers with the highest effect on researchers. It concentrates heavily on redemptive overall impact data. Whether a scientist has a low impact career and is strongly focused on the research field and the duration of his entire career. The measuring system based on the quantity of references does not accurately evaluate its importance.

Fukuzawa et al. have already been studying the dissemination of patented research publications and citations. The research also contained an evaluation of their friendship. The research comprised of more than 4,000 publications, and Japan's leading authors were considered. The study showed the relationship between them in U-shape [34].

Jin et al. Suggested method by which h-index and r-index are combined. The h-index and the ar-index are also combined. Such coupled-indices were used to

establish the criteria for the researcher 's assessment. The pair, h-Index and Ar-Index have proven to be a better predictor than the other indicators used to classify individual researchers. Schreiber et al have analyzed h-index and its 17 variants, which can be classified as, quantitative and qualitative parameters. A meta-analysis of h-index variants was carried-out as well. This meta-analysis showed a large and easy association between the h-Index and its variants. It also has shown that, these variants have not added much to the original h-index [35]. In 2006, the combination of h-index and g-index was discussed. This relationship has been confirmed by data from 19 chemical professors at the University of Poland. The experiment showed a clear association between h-index and g-index.

The h-index was compared to several other h-variants in the Van Rann efficiency test. Instead of a particular sample, this analysis was aimed at defining the most active research group. Only take the example of the last 3 years into account. Van Rann uses data from 147 chemistry study groups in the Netherlands to determine the similarity between variants of h-index to test the scientific results [36].

In 2016, Ayaz and Afzal Presented a new method of testing h-variants, recognizing the champions of the scientific community in the field of mathematics. The production of h-variants was assessed by counting the number of occurrences in the index column. As a result, full-h was higher than g-index and h-index [10]. In another study, h-index, g-index, q2-index etc. have been evaluated on mathematics dataset by considering the same bench mark dataset used by Ayaz and Afzal [10]. They have concluded that, h-index performed better than the g-index, q2-index, hg-index, R-index and A-index, for elevating the awardees at top of the list. However, h-index only managed to carry 33 awardees in the list 5,753 authors [16].

In the field of neuroscience, in 2019 Ameer and Afzal examined the h-index and its quantitative and qualitative variables such as g-index, hg-index, m-quotient, e-index, f-index and r-index respectively. Awards winners for measuring the success of all such metrics from scientific neuroscience groups called benchmarks, was used. The above findings showed that the estimates of the hg-index and R-index were higher than the other indices, both of which increased 25 awardees in the list of 55 authors [17]. Critical analysis of the state-of-the-art literature is shown in

Table 2.6.

The intensive literature, number of metrics used to determine and rate the value of the publications or to classify the most convincing scholars have been studied. The derivations that we have found are, Scientists nearly 10 years earlier, by the help of their publications and citations, were graded. Later in the beginning of the decade, without taking into account the shortcomings and background of study, the process model began using variants of the H-Index. Once a new methodology is commonly introduced in this domain, it is typically developed unconventionally or with various data sets. Since these techniques are based on different data sets and validated, it is difficult to perceive the importance of these techniques individually.

In order to rank these parameters, a standard method is required. According to which these criteria could be used to introduce a new index that could essentially rate researchers and study groups, etc.

TABLE 2.6: Critical analysis of Literature review

Ref	Dataset	Indices	Result	Limitation
[13]	**Ingenious example:** **scientist X1:** publications ->20 Citations: One paper cited->10 All other ->1 **scientist X2:** publications->30 Citations: One paper cited ->10 All other->2	A-index, R-index.	For X1: A-index->10 R-index->3.16 For X2: A-index ->6 R-index ->3.46	Proposed for imaginary scenario.
[16]	**Field:** Mathematics **Authors:** 57,533 **Publications:** 69,197	h-index, A-index, q^2-index,	Elevate 30.86% of awardees in top 10% using	Domain specific dataset

Ref	Dataset	Indices	Result	Limitations
	Citations: 8,821,251 **Awardees:** 461	g-index, R-index, hg-index.	h-index	
[17]	**Field:** Neuroscience **Authors:** 546 **Publications:** 96,317 **Citations:** 5,850,906 **Awardees:** 48	h-index, m-quotient, e-index, A-index, hg-index, f-index. R-index	Elevate 52% of awardees in top 10% using h index	Domain specific dataset
[37]	**Field:** Civil **Field:** Engineering **Authors:** 36,921 **Publications:** 20,307 **Citations:** 2,184,638 **Awardees:** 1060	h-index, t-index, q^2-index, hc-index, A-index, f-index, wu-index.	Elevate 47% of awardees in top 10% using f-index	Domain specific dataset
[10]	**Field:** Mathematics **Authors:** 57,533 **Publications:** 69,197 **Citations:** 8,821,251 **Awardees:** 461	h-index, hc-index, g-index.	Elevated 95 awardees in top 1000 using hc-index	Domain specific dataset
[38]	**Universities:** 39 **Authors:** 57,533 **Publications:** 58,781	h-index.	Ranked central universities in India	No evaluation is performed

Ref	Dataset	Indices	Result	Limitations
[39]	**Field:** Neuroscience 37 neurosurgical departments'data	h-index between (2004-2014) and (2012-2014)	Ranked neurosurgical departments	No evaluation is performed

2.5 International Awards

Since no standard benchmark dataset is available in this area of study. Dunaiski et al. used the award winners to evaluate such indices for the first time [40]. Later on, Ayaz and Afzal presented a novel technique by using prestigious science societies' award winners as benchmarks [10]. Later on, the same method is used by Raheel et al [37], Ain et al [16], and Madiha and Afzal [17]. However, we considered the recipient of the award not only as a benchmark in this research but also for evaluation. Exceptional researchers are given awards by their respective scientific societies. The research is focused on researchers in three fields, Civil Engineering, Mathematics and Neuroscience. For each of the above fields, we have selected 250 award winners. For Civil Engineering, award recipients belong to following awarding societies: i) ACI, ii) ASCE, iii) CSCE, and iv) ICE. Considered awarding societies from Mathematics are, i) IMU, ii) LMS, and iii) AMS, and Neuroscience societies are, i) SFN, ii) FENS, iii) CNS, and iv) ANS. In next section, we discussed above mentioned societies and their effectiveness.

2.5.1 Civil Engineering Societies

(i) **American Concrete Institute**

The American Concrete Institute is a non-profit research organization and a body of developing standards. The ACI was set up in January 1905. ACI is currently located in Farmington Hills, Michigan, USA. ACI's goal is to develop and encourage a consensus on cognitive implementation.

(ii) **American Society for Civil Engineering**

The American Civil Engineering Society was established in 1852 as a tax exemption for the Civil Engineering industry worldwide. It is the oldest major US electronics firm in Reston, Virginia. This was founded in 1848 by the old Society of Boston Civil Engineering. The ASCE is committed by way of civic leaders' efforts to promote civil engineering study and education and the advancement of human rights. It has about 152,000 members in around 177 nations. This aims to provide valuable public interest, to promote the advancement of technologies and to build and sustain civil engineers.

(iii) **Canadian Society for Civil Engineering**

The Civil Engineering Society of Canada was founded in 1887, re-named the Engineering Institute of the Canadian Society of Civil Engineering in 1918 and reestablished as a member of the EIC in June 1972 [1]. This promotes civil engineering advancements including geotechnical, electrical, technological, hydrotechnical, geological, transport and surveying and geomagnetic, etc. Civil Engineering members are typically classed as members and are qualified to be nominal for an Associate Member or fellow Members of the Registered Civil Engineering Group. There are also fields of research at other institutions across Canada, such as University of Toronto, McGill, and the Canadian Association for Civil Engineering University of British Columbia. Osama Moselhi received its best design paper in 2019 from the CSCE. During the conference on construction specialties the Best Paper Award is awarded every two years.

[1] https://csce.ca/en/

(iv) **Institution of Civil Engineers**

The Civic Engineers Institution (CEI) was founded 202 years ago in 1818 and is an independent professional civil engineering organization in the UK. In London, ICE employs over 92,000 staff, three-quarters of whom are in the United Kingdom and the rest are in 150 foreign countries. ICE aims to encourage the field of civil engineering through technical qualifications, advancement of knowledge, professional and ethical growth and business, academia and government contact. It provides training, recruiting, social service and logistics services for the corporate sector. As an institution, ICE seeks to facilitate and support planning, technical ethics administration and activism in the area of engineering. It sets standards for the organization; it establishes technical practices for industry and academics and provides instruction and program guidance.

2.5.2 Mathematics Societies

(i) **International Mathematical Union**

IMU is a research body to facilitate global cooperation in mathematics [2]. The Committee would therefore aim to organize conferences and lectures in all fields of mathematics.

(ii) **London Mathematical Society**

LMS is based in the UK and offers mathematics and research publications [3]. It also provides students with scholarships and offers workshops and seminars.

(iii) **American Mathematical Society**

The American Mathematical Society was established for mathematical research and grants [4]. The state and the international mathematical world greatly benefit from their talks, conferences and many other events. Through

[2]https://www.mathunion.org/
[3]https://www.lms.ac.uk/
[4]https://www.ams.org/home/page

the assistance of many other organizations, the American Mathematics Society organizes the highest national professional gatherings. Publications include scientific journals, textbooks and articles on research.

2.5.3 Neuroscience Societies

(i) **Society for Neuroscience**

The neuroscience society was founded in 1969 as a professional society. Currently, it has over 37,000 members and its offices are devoted to fundamental doctors and researchers from around the world in Washington, DC [5]. This is particularly well known for its annual convention as one of the world's largest research gatherings.

(ii) **Federation of European Neuroscience Societies**

The European Neuroscience Society Federation was founded in 1998 for organizing and sharing research on global neuroscience systems both nationally and globally. This substituted for the American Neuroscience Society. In total, FENS is made up of 42 collaborations close to 23,000 researchers [6].

(iii) **Cognitive Neuroscience Society**

The Cognitive Neuroscience Society is an international research Society with an interest in the cognitive function of the brain [7]. Participants in society engage primarily in cognitive neuroscience to study the integration of our brain and mind experience, which is focused on biologically derived and psychological evidence. Society comprises of scientific groups from several fields, which was founded in 1994 by a group of six scientists: Michael S. Gazzaniga (University of California, Santa Barbara). The Society includes George R. Mangun (UCD), Steve Pinker (University of Harvard), Patti Reuter Lorenz (UCD), Daniel Schakter (UCH) and Art Shimamura (University of California, Berkeley). The Cognitive Neuroscience Society publishes the Cognitive Neuroscience Journal in the MIT Press.

[5] https://www.sfn.org/
[6] https://www.fens.org/
[7] https://www.cogneurosociety.org/

(iv) **Australian Neuroscience Society**

The Australian Neuroscience Society was founded as an informal group of interested Australian Neuroscientists in 1971 [8]. Around 1972 and 1980 annual conferences on a specific subject were held. At the 1979 "Neurotoxins" conference in Flinders, it was determined to create an effective tradition. During the Canberra conference in 1980 a formal society was founded and the Council sat for the first time on Thursday, 7 February 1980.

[8] https://www.ans.org.au/

Chapter 3

Research Methodology

The scientific community continues to build a number of methods for categorizing researchers in their respective fields. The results of Chapter 2 indicate that researchers are typically classified using publications, citations, h-index and some of their combinations. However, no systematic research has been found to identify these writers by rating parameters based on their effectiveness. The evaluation of the proposed index is based on distinguished national and foreign awardees and non-awardees from the fields of Civil Engineering, Mathematics and Neuroscience. This section describes the methodology while the scheme of the proposed technique is illustrated in Fig. 3.1.

3.1 Domain Selection

Comprehensive data from diversified academic fields are required to perform experiments. We have chosen the field of Civil Engineering, Mathematics and Neuroscience to derive the weights of basic parameters and h-variants according to their usefulness for the recipients in each class. Such fields have been chosen, after careful consideration of the various factors listed below, with the respective field section.

3.1.1 Civil Engineering

The Civil Engineering domain is one of the oldest areas of human life and significant research work is carried out in this area, which is why the Civil Engineering domain is chosen to test the proposed methodology. In addition, each year, according to the contribution of their work, there are many scientific societies evaluating the distinctive researchers in the field. However, researchers in this field have not been used extensively for the assessment of h-indices. The weights of the parameters of the research assessment help the scientific community in this field to assess the merits of the individual and to promote the development and growth of this area. Therefore, due to its significance in the modern age of building, this area should be further explored.

3.1.2 Mathematics

The Mathematics Domain provides the complexity required to check the proposed index. The connection between mathematics and all other fields, including IT, engineering and chemical, etc., is another significant explanation for selecting this field.

3.1.3 Neuroscience

We have selected the Neuroscience domain for the evaluation of the h-index, its variants and the proposed index. The reason why the field of Neuroscience is chosen is that significant research is being done in this area. Neuroscience studies the symptoms of the illness and the wound regions of the nervous system. Within this area, scientists are working tirelessly to find ways to prevent or treat brain, nervous system, and body complications. More than 1,000 brain and nervous system disorders are present. The classification of Neuroscience researchers is therefore a necessary concern for the recognition of a qualitative field scientist.

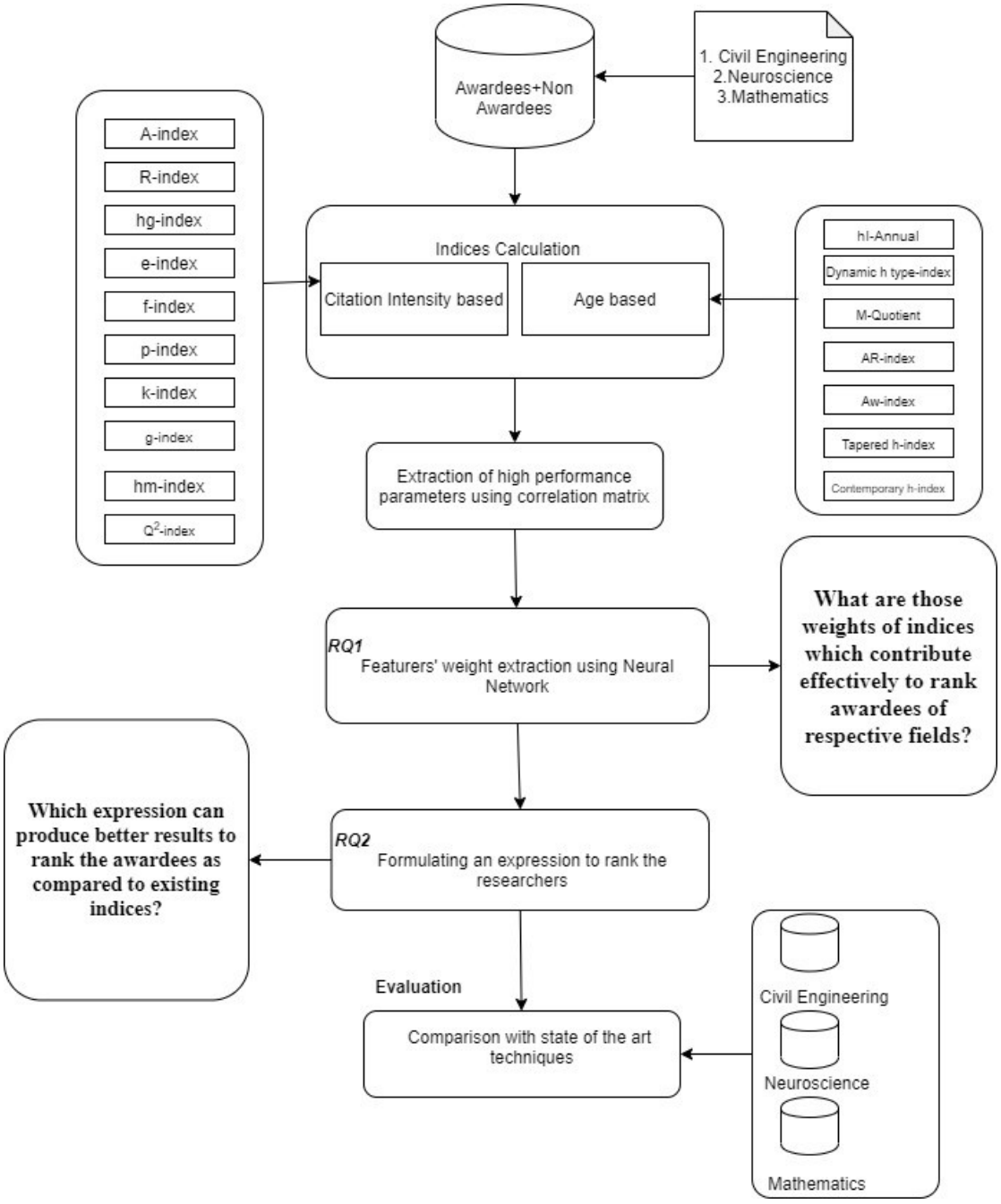

FIGURE 3.1: Methodology Diagram

3.2 Dataset Acquisition

Civil Engineering is one of the oldest fields of research dedicated to the creation, construction, maintenance of natural and physical environments such as dams, airports, sewerage, structural building components, etc. The dataset of Ayaz and Afzal was considered to collect the list of non-recipients [10]. They used the certified Civil Engineering Database (CEDB), a qualified initiator of the ASCE. ACSE (American Society of Civil Engineering) is a renowned civil engineering research society. In addition, 250 awardees of prestigious Civil Engineering scientific societies (ACI, ASCE, CSCE, and ICE) from 2011-2019 were considered for assessment.
For the collection of non-award recipient lists for the Neuroscience domain, Ameer and Afzal datasets were considered. They used the Neuroscience field theme classification to collect the data. The NIF (Neuroscience Information Framework), a synthesis of the NIH Blueprint for Neuroscience Research, is an approved Neuroscience classification. Renowned Neuroscience research societies such as SFN, FENS, CNS and ANS have been considered for awardees of a total of 250 Neuroscience awards [17].
The non-awardees list belonging to the Mathematics domain is based on the Ain et al dataset [16]. They used a crawler, which is built by Ayaz and Afzal [10]. Ain et al. uses the term Mathematics subject classification to crawl Google Scholar's results. Prestigious research organizations such as IMU, AMS and LMS have been selected for the list of 250 recipients of the award.

3.3 Search Engine

This research used Google Scholar to collect information regarding authors, including publications and citations from a perfectly tailored list of award winners and non-awardees as shown in Fig.3.2 and Fig. 3.3.
There are various sources of information about authors from each discipline is available like Web of Science (WOS), Google scholars, Scopus, etc. Researchers

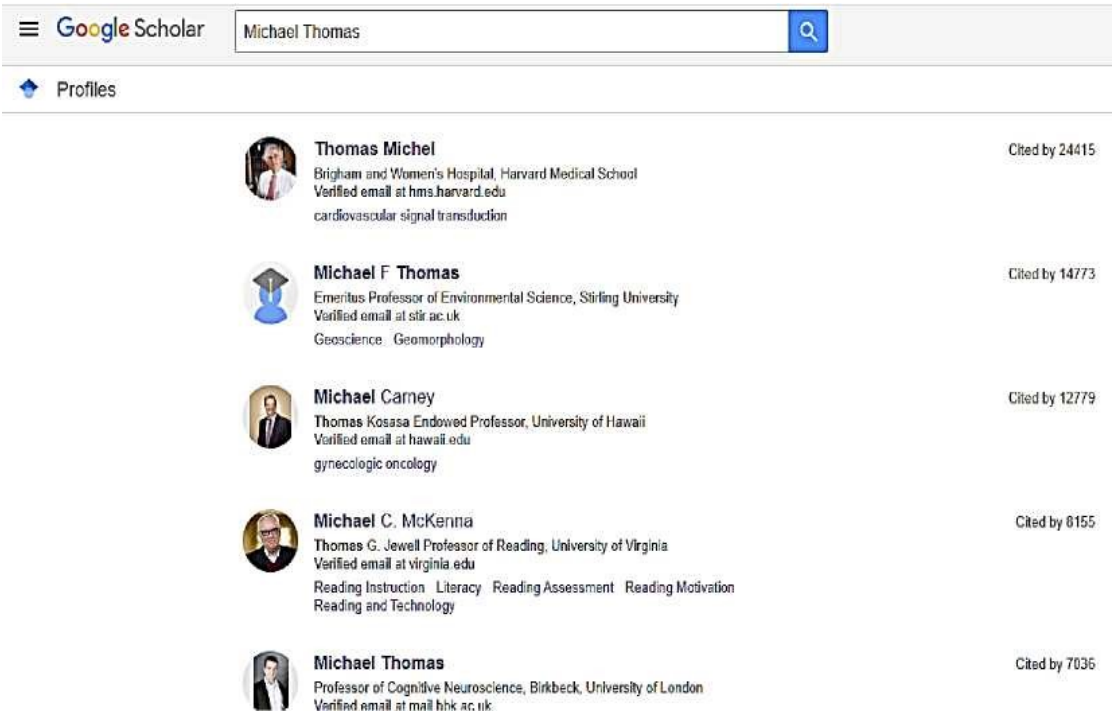

FIGURE 3.2: Search Engine

have free access and limited reach issues for all services except for Google Scholar [10]. As Google scholarly data is freely accessible and encompasses all fields, we have selected Google scholars rather than Scopus, WoS and other sources. There is sufficient research carried out to compare Google scholars with the Web of Science (WOS) data. Harzing also reported that Google scholars are better than WoS, as new data is updated regularly [41]. We have started collection of data in Aug, 2019.

3.4 Data Cleaning

All publications were gathered from Google scholar considering the top ranked relevant and highly cited results obtained by querying the targeted keywords given by domain experts on Google scholar; these publications have undergone two tests to ensure their relevancy with the domain and to filter duplicate occurrences of authors in the gathered data.

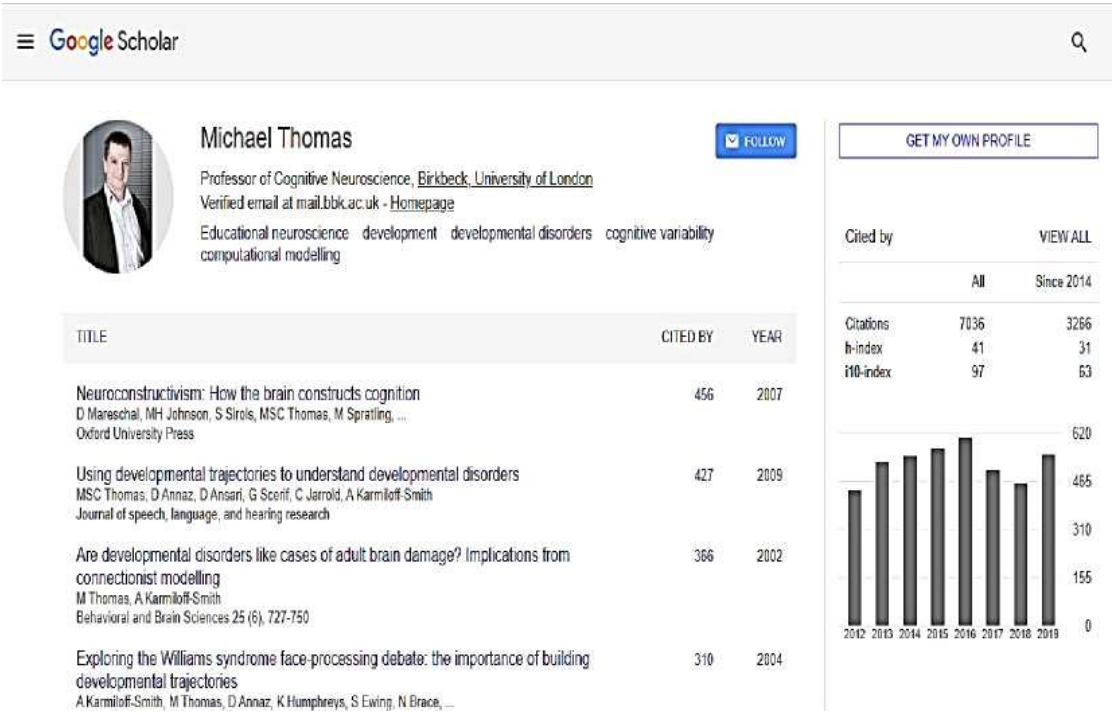

FIGURE 3.3: Search Engine

3.4.1 Relevance with the Domain

(i) Those publications were removed which were not actually the publications and their titles have special symbols such as: '?' , '*', and '#' etc.

(ii) A random sample was given to the domain experts for determining its relevancy with the field.

(iii) Then the publications were filtered on the basis of their publication venues and publications that were not related to domain were discarded.

3.4.2 Author Disambiguity

There can be two scenarios of name variations. The first one is that the last names and first names both are same i.e. we need to verify whether these two authors are actually different or identical. The second one is that the last names are same but first names are different. There are more chances that both will be different but still need to be evaluated.

(i) For first scenario each publication of the authors stored in the database were checked to ensure that they belong to the same author or different. The result of this process discloses that no two authors exist in the database which shared the same last and first names.

(ii) For second scenario, every authors' homepages are visited and publications are matched with those stored in the database. Most of the authors were those whose last names were same and first names were different and they were in reality different authors while some authors had the same last names and different first names, but they were name variants of one author.

3.5 Dataset Description

Finally, there are three databases from three diversified fields, Civil Engineering, Mathematics and Neuroscience. The dataset of each domain includes the data of 500 researchers. Among the 500 scholars, 250 are recipients of awards from prestigious scientific societies in the respective field and 250 are non-recipients. In the field of Civil Engineering, award winners are from ACI, ASCE, CSCE and ICE. The award winners for Mathematics are from IMU, LMS and AMS. Neuroscience domain award recipients are from SFN, IBRO, FENS, CNS, and ANS. The detail description of dataset is shown in Table 3.1

Table 3.1: Dataset Description

Dataset Description	
Total No of Authors	1500
Total No of Publications	376,963
Total No of Citations	68,049,621
Civil Engineering	
Awardees Description	
Total No of Authors	250

Total No of Publications		46,465	
Total No of Citations		1,770,447	
Non Awardees Description			
Total No of Authors		250	
Total No of Publications		37,730	
Total No of Citations		2,306,275	
Awarding Societies Description			
Name	**Found in**	**Members**	**Award Winners**
ASCE	1852	152,000	158
ACI	1904	30,000	14
CSCE	1887	75,000	64
ICE	1818	90,000	14
Mathematics			
Awardees Description			
Total No of Authors		250	
Total No of Publications		82,488	
Total No of Citations		49,163,332	
Non Awardees Description			
Total No of Authors		250	
Total No of Publications		60,117	
Total No of Citations		3,245,309	
Continue on Next Page			

Table 3.1 – continued from previous page

Awarding Societies Description			
Name	**Found in**	**Members**	**Award Winners**
AMS	1888	30,000	92
LMS	1865	3,000	96
IMU	1920	27,000	62
Neuroscience			
Awardees Description			

Total No of Authors	250		
Total No of Publications	76,712		
Total No of Citations	7,418,106		
Non Awardees Description			
Total No of Authors	250		
Total No of Publications	73,451		
Total No of Citations	4,146,152		
Awarding Societies Description			
Name	**Found in**	**Members**	**Award Winners**
SFN	1969	37,000	79
IBRO	1961	41,000	28
FENS	1998	23,000	31
CNS	1994	2,000	37
ANS	1971	19,000	74

3.6 Tools and Technologies Used

(i) PyCharm (for python)

(ii) Publish or Perish

3.7 Selection of Author Assessment Parameters

The data obtained from the Google Scholar is stored in a CSV file (comma separated) using publish or perish tool for the list of non-rewarded researchers and award recipient researchers. After calculation of the indices, we evaluated each of the primitive parameters, citation-intensity and age-based parameters using the correlation matrix.

3.7.1 At the end of this step, we will be able to come up with the most efficient author evaluation parameter for each category and best among all inclusive primitive parameters (papers, citations, years, Authors / Paper, Citations / year, Citations / Paper, H-core-citation, and h-index), Citation-based parameters (g-index, hg-index, A-index, R-index, f-index, p-index, q2-index, k-index, and e-index), and age-based-parameters (m-quotient, hc-index, hl-index, hl-norm, Aw- index, hl-annual, and Ar-index).

3.7.1 h-index

In 2005, Jorge Hirsch, a respected German physicist, proposed the h-index [9]. It is a metric used to calculate the influence of authors by taking their publications into account. The papers of the author are given in descending order with respect to the citations of C(i). If a scientist's citations and publication evidence are accurately skewed, the H-Index does not shows a complete picture.
Take example of 2 researchers, A and B; A has a few articles but all are highly cited. While B has a large number of publications, they all have low number of citations. The h-index is the same for both. Imagine another author, whose h publications are each cited in h times, but this is an unrealistic case, because such discrete outputs are not relatively common. The h-index indicates the incomplete representation of researchers who've had diversified publications and citations.

3.7.2 g-index

The extension of the h-index, called the G-index, is introduced as a measure of the author's impact on the basis of the article quotations. The g-index is an index expected as an expansion of the h-index According to Egghe et al. g-index calculates the importance of best articles by authors and it is categorized as: "As the highest number of g papers receiving together g2 or more citations of this type, it is already obvious that g¿=h" [11]. The distinction between the two is that the g-index gives higher priority to a paper with a larger number of citations. This

would benefit to credit articles with less quotations. It is going to change. G-index has a broad range of scenarios and fits well in researchers' rankings.

3.7.3 A-index

The effect of a publication can be evaluated in the total citations, the h- index examines the most enormous influence of the author's work, which named as Hirsh Core by Hirsh himself in 2007. Although, In the year 2007, this was Burrell's claim [45]. In addition, In Jin et al [46], Hirsh Core was described more explicitly as the highest and greatest published paper set. In accordance with the core works of the Hirsch, the A-index is described by the maximum number of citations. As suggested in 2006, it should be treated as an H-Index variant since it deals with the most commonly cited papers [12]. A-index is defined as:

$$A - index = \frac{1}{h} \sum_{j=1}^{h} Cit_j \tag{3.1}$$

3.7.4 R-index

A-index gives the lowest rating to high-h-index researchers since it uses the h-index as a divisor. The more h-index the researcher gets, the less A-index he has. Investigators thus concluded that improved results can be obtained by taking the square root of the cumulative number of citations instead of dividing by h. R-Index only counts the number of citations in h-core, such as A-index, so that it may be very vulnerable to just a few heavily cited papers. Jin et al. said that the A-index separates the h-index by A-index, thereby inherently biased towards authors with large h-index values [13]. They proposed that the square root of the citations in h-core should be calculated instead of dividing by h, removing the bias problem. The R-index estimation method is much like A-index, and only a limited number of articles with a lot of citations will easily affect it. The equation

of R-index is:

$$R - index = \sqrt{\sum_{j=1}^{h} Cit_j} \tag{3.2}$$

3.7.5 hg-index

In 2006, Rousseau suggested that several aspects of the author's work were determined by g-index and h-index. The argument, however, is that none of them can provide a full picture, including all facets. S. Alonso et al. noted that all indices would be taken into consideration in order to produce quantitative performance, all of which rely on measuring the various features of the research papers. This idea leads to a proposal for a new hg-index [14]. The hg-index eliminates the relative weaknesses of the 2 indices and incorporates their strengths. Hg-index can be calculated by taking geometric mean of product of both g-index and h-index.

$$hg - index = \sqrt{h * g} \tag{3.3}$$

3.7.6 f-index

The f-index represents the fold of h2 citations received for papers in h core [42]. It is a fractional citation counting scheme. Mathematically it is represented as follows:

$$f - index = (\frac{e}{h})^2 \tag{3.4}$$

3.7.7 P-index

The h-index has quickly grasped the picture of Scientometricians and Bibliometricians. This was intended to be an easy way to calculate combined measure that wanted to be incorporated into a single number. The h-index, however, applies to the citation and publication without defining the connection [43]. Whereas, p-index is the best balance between the total citations and the average citation

rate for publication. The p-index is more robust than the h-index as it offers the optimal balance of quantity and consistency [26]. Mathematically it is represented as follows:

$$p - index = (\frac{c^2}{h})^{\frac{1}{3}} \tag{3.5}$$

3.7.8 q^2-index

The q^2-index is a combination of the h-index and the m-index. Quantitative calculation is provided by the h-index, while the m-index is a qualitative measure. The h-index is of a quantitative nature and gives insight into the number of articles. The h-index is cautious and gives an extensive number of papers, whereas the m-index is a qualitative metric which deals appropriately with the distributions of the citation volume. The q^2-index takes into consideration both the qualitative and the quantitative dimensions. It is also safe to assume that this metric has a qualitative and quantitative component that gives a fuller view than when one item is considered alone. The q^2-index is not minimally impacted by other high values and uses geometric mean to simplify the index interpretation [24].

$$q^2 - index = \sqrt{h * m} \tag{3.6}$$

3.7.9 K-index

The h-index consists of two collections of citations for reference. Set of citations for h-core publications and h-tail publications. H-index just took into account the quotes in h-core and skipped the quotes in h-tail. The growth in h-core citations does not lead to the work impact of scholars, except in the case of h-core citations. Ye and Rousseau also developed a new index that also takes into account citations in h-tail, dubbed k-index [44]. We found that the tail-core ratio usually increases in actual cases, while the k-index declines. This has been shown that the principle of power-law is compatible with certain practical results. Mathematically it is

represented as:

$$k - index = \frac{\left(\frac{c}{p}\right)}{\frac{c(h-tail)}{c(h-core)}} \tag{3.7}$$

3.7.10 e-index

Improved h-core quotes do not add to the effect of researchers on the h-index situation. The e-Index also takes into account the neglected h-core citations. Compared to some other h-index metrics, the E-index has some weaknesses in the calculation. Increased citations in h-core do not contribute to the impact of researchers in h-index. The e-index also takes into account missed citations in the h-core. Compared to several other indices that rely on the h-index, the e-index still has certain drawbacks related to its calculate. Dependency means that the e-index cannot be calculated without the value of the h-index. Citations of h-tail articles have therefore not been taken into account by the e-index. In the event that all publications have citations greater than h-index, the e-index will be zero and the h-index will not lose any information. However, such a case never occurs. It's very hard to find a researcher with such good results. It is clear that publications need some time to make citations [23]. Mathematically it is represented as:

$$e - index = \sqrt{d^2 - h^2} \tag{3.8}$$

3.7.11 m-quotient

Some of the primary drawbacks of the h-index is that researchers are still optimistic, even when field work is not involved while the influence of the research remains in progress. This is also not possible to equate researchers of varying ages with the h-index, given that the h-index is weighted against older researchers. As the h-index and the length of the career are roughly directly proportional to each other. To resolve this h-index deficit, Hirsh himself suggested m-quotient. Taking into account the question of career time, Hirsch put forth the m-quotient or the m-quotient of Hirsch in his original work. Since the first publication, he has split the

h-index to equate researchers of various career lengths [45]. The m-quotient is also helpful because it is appropriate to compare researchers of different career lengths. The key drawback of the m-quotient is that if a writer fails to contribute after some time, the h-index continues to decrease. Mathematically it is represented as:

$$m - quotient = \frac{h - index}{y} \tag{3.9}$$

3.7.12 hI-norm

Instead of considering only citations in h-core, such as h-index, hl-norm took standardized citations into account. It normalizes the number of citations for each article, and then calculates h-Index for standardized references by dividing citations by author number for each article [32]. This is much more advanced technique than that. As normalized quotes which were not included in the original h-index were taken into account.

3.7.13 hc-index

The h-index does not take into account the age of the paper as the writers who received a higher h-index from a big publication did not receive a deduction for retired or inactive research work. Sidiropoulos et al, regarded as the contemporary h-index [28], have suggested a simplified version of h-index. It is skewed and pays more attention to the recently released articles. This shall assess the effect of article i, which is referred to as the hc-index. This simplified index takes quotations but gives more attention to the new article.

$$hc - index = \gamma * \frac{C(i)}{(Y(now) - Y(i) + 1)} \tag{3.10}$$

Where Y(now) reflects the current year, Y(i) is the year of the publication, and C(i) refers to the citations of paper i. As the number of citations is divided with times, so the value of hc-index is so small. The γ factor comes in to deal with this

deficiency. This can regularly reduce the impact of article with the passage of four year of time interval as shown in below equation.

$$hc - index = \frac{C(i)}{1}, \frac{C(i)}{2}, \frac{C(i)}{3}.....\frac{C(i)}{n} \tag{3.11}$$

3.7.14 Aw-index

The Aw-index is influenced by the Ar-index and is known to be the age of each publication. We did, however, find all the articles written by the scholar instead of including only highly-quoted articles. The Aw-index is defined as the amount of the average number of citations of all articles each year [31]. A successful policy would be resistant to transition over time. This index will indicate a performance increase and time decrease. The Aw-Index is the square root of the total average annual citations for all publications. Mathematically it is represented as:

$$Aw - index = \sqrt{\sum_{j=1}^{n} \frac{Cit_j}{a_j}} \tag{3.12}$$

3.7.15 hI-annual

A study of researchers with different career periods has been dealt with, suggested hl-annual. It refers to the annual average change in the individual h-index [30]. Eventually, since the metric reflects the total number of independent publications written annually by a researcher, an empirical definition is often possible.

3.7.16 Ar-index

Not only did the Ar-index find the citations in h-core, it also found the rise in citations in h-core, but also the age of article. The Ar-index shall be specified as the amount of the average number of citations per year of the articles included in the h-core [13]. According to Jin et al., a successful method would be resistant

to change over time. This measure has the potential to display a rise and decline in results over time. The square root of the cumulative annual average quotes for all articles in the h-core is the Ar-Index. The Ar-Index has three factors: the publication of their citations and the number of years after the first publication. Ar-index is meant to eliminate stigma against writers who have not been involved in work for some time since the h-index does not decrease with times of inactivity.

$$Ar - index = \sqrt{\sum_{j=1}^{h} \frac{Cit_j}{a_j}} \tag{3.13}$$

3.8 Correlation

Correlation is a statistical method used to calculate the association between two or more variables, i.e. the extent to which the variables are related, so that the changes in one of the variables are followed by a change in the other. The correlation is said to be linear if there is a constant ratio of change in the quantity of a variable to that of another. In comparison, if the proportion of the change in one variable to that of a change in another variable is not constant, the correlation is non-linear or curvilinear.

(i) Scatter Diagram Method

(ii) Karl Pearson's Coefficient of Correlation

(iii) Spearman's Rank Correlation Coefficient; and

(iv) Methods of Least Squares

Above are the methods to measure the relationship between two or more features. We have used scatter diagram method in our research since scatter diagram method is based on the study of graphs while the rest is mathematical methods that use formulae to calculate the degree of correlation between the features.

3.9 Parameter Weights

There are different machine learning and deep learning techniques that can be used for parameter weight extraction i.e. i) Linear regression ii) LSTM iii) Artificial neural network iv) Convolutional neural network and other variants of neural network. To answer the first research question, we have ranked the quantitative author ranking parameters based on their importance by using neural network. We have selected this technique because during the training of a network, the same set of data is processed many times as the connection weights are continually refined [46].

We have passed all indices as input to the input layer and hidden layer is having three nodes and one output node. Binary sigmoid activation function is used in our experiment since our label is either one or zero. We have trained the network on fifty epochs since increasing number of epochs was not increasing the prediction accuracy. After fifty epochs we were having maximum accuracy of this network to predict awardees and weights of all parameters were extracted to reach

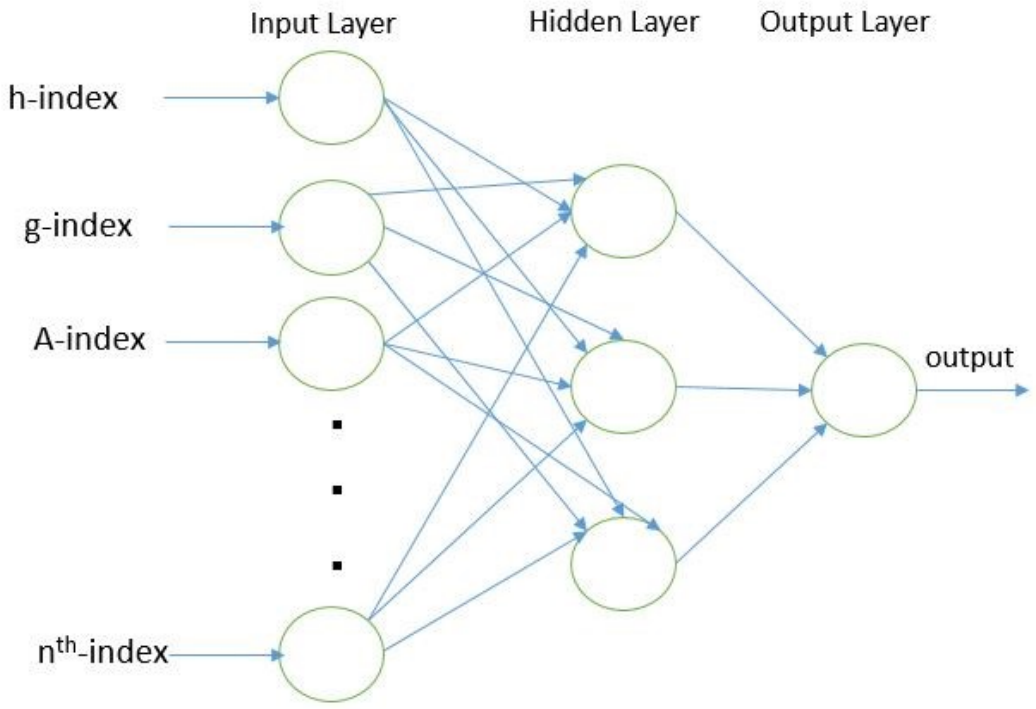

FIGURE 3.4: Neural Network

TABLE 3.2: Neural Network weights

Features	**Weights**
Authors/paper	0.75
f-index	0.61
Years	0.29
K-index	0.17
HI-annual	0.12

to this maximum accuracy. These weights shows relationship with the label i.e parameters having higher weights shows it is bringing more awardees as compare to other indices.

3.10 Index Formulation

After finding correlation between features i.e. primitive, citation-intensity and age-based parameters based on their effectiveness for the award recipients of Civil Engineering, Mathematics and Neuroscience, individually as well as in combination. We have found the effectiveness of those features using neural network and the top 2 features are used based on their weights to formulate a new index. We have ranked the researchers according to ST_{index} from each field, and analyzed the trend of award recipients in the list of top 100 researchers.

Chapter 4

Result and Evaluation

This chapter describes the findings obtained by each step of the methodology mentioned in Chapter 3, which was introduced to formulate an index to rank awardees in the fields of Civil Engineering, Mathematics and Neuroscience.

4.1 Correlation of Assessment Parameters

To answer the first research question, in this section, we have found correlation of primitive, citation-intensity and age-based parameters to analyze which author assessment parameters are bringing unique awardees at top ranking. Secondly, weights of top 5 parameters are extracted using neural network to identify the best author assessment parameters for award recipients among the all these parameters. After combining impact of the top two parameters, it is compared with the baseline papers and has elevated 85% of awardees for Civil Engineering. After combining impact of the top two parameters, it is compared with the baseline papers and has elevated 85% of awardees for Neuroscience domain. After combining impact of the top two parameters, it is compared with the baseline papers and has elevated 83% for Mathematics domain. When we have combined all the researchers from Civil Engineering, Mathematics and Neuroscience domain, the proposed index has achieved 80% results.

TABLE 4.1: Weights Civil

Features	weights
Author/paper	0.71
A-index	0.52
Years	0.21
F-index	0.17
K-index	0.15

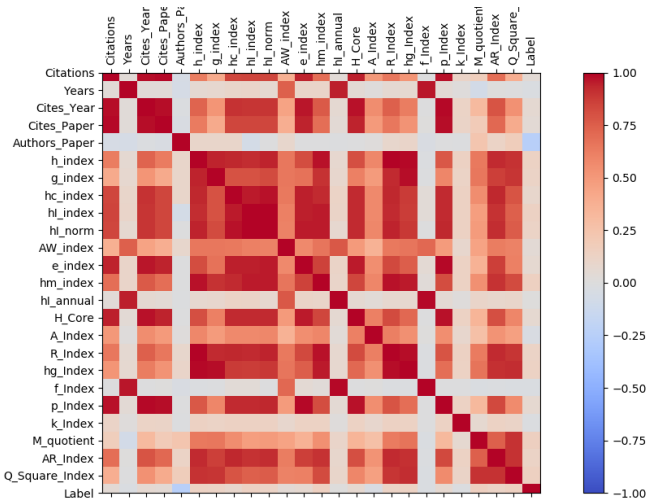

FIGURE 4.1: Correlation Civil

4.1.1 Civil Engineering

The top five parameters for the Civil Engineering, i) Authors/Paper, ii) A-index, iii) Years, iv) K-index, and v) f-index are extracted using correlation matrix shown in Figure 4.1. We can see indices along x-axis and y-axis and color scheme shows relationship among these parameters. Weak relationship among two parameters means that they are bringing unique results as compare to each other. On the basis of weak relationships among each other, these five parameters are selected. Neural network back propagation approach having three hidden layers is applied on these 5 parameters to find out weights of these parameters as shown in Table 4.1. Based on the weights having the maximum impact to rank awardees, top two parameters Authors/paper and A-index are selected to formulate the index. Authors/paper

and A-index had strong relation with the label and weak relation with each other, means they are bringing maximum number of Awardees but different from each other. Hence, $ST_{c(index)}$ for Civil Engineering domain is as follows:

$ST_{c(index)} = A_{index} + 1/(Authors/paper)$

4.1.2 Mathematics

The top five parameters for the Mathematics domain, i) Authors/Paper, ii) f-index, iii) Years, iv) M-Quotient, and v) A-index are extracted using correlation matrix shown in Figure4.2. We can see indices along x-axis and y-axis and color scheme shows relationship among these parameters. Weak relationship among two parameters means that they are bringing unique results as compare to each other. On the basis of weak relationships among each other, these five parameters are selected. Neural network back propagation approach having three hidden layers is applied on these 5 parameters to find out weights of these parameters as shown in Table 4.2. Based on the weights having the maximum impact to rank awardees, top two parameters Authors/paper and f-index are selected to formulate the index. Authors/paper and f-index had strong relation with the label and weak relation with each other, means they are bringing maximum number of Awardees but different from each other. Hence, $ST_{m(index)}$ for Mathematics domain is as follows:

$ST_{m(index)} = f_{index} + 1/(Authors/paper)$

TABLE 4.2: Weights Maths

Features	weights
Author/paper	0.87
f-index	0.70
Years	0.25
a-index	0.21
m-qoutient	0.19

4.1.3 Neuroscience

The top five parameters for the Neuroscience, i) Authors/Paper, ii) K-index, iii) A-index, iv) f-index, and v) Years are extracted using correlation matrix shown in Figure 4.3. We can see indices along x-axis and y-axis and color scheme shows relationship among these parameters. Weak relationship among two parameters means that they are bringing unique results as compare to each other. On the basis of weak relationships among each other, these five parameters are selected. Neural network back propagation approach having three hidden layers is applied on these 5 parameters to find out weights of these parameters as shown in Table 4.3. Based on the weights having the maximum impact to rank awardees, top two parameters K-index and A-index are selected to formulate the index. K-index

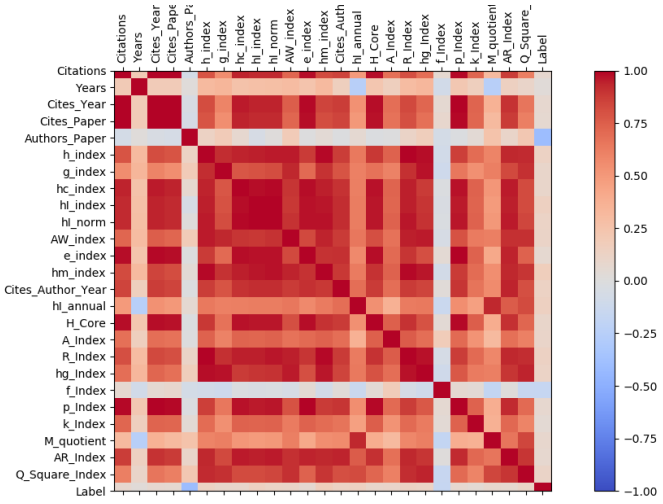

FIGURE 4.2: Correlation Math

TABLE 4.3: Weights Neuroscience

Features	weights
A-index	0.54
k-index	0.53
Author/paper	0.51
f-index	0.34
years	0.17

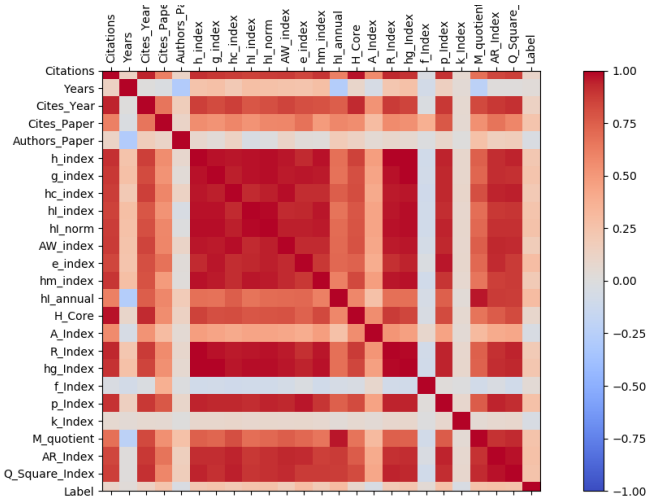

FIGURE 4.3: Correlation Neuroscience

and A-index had strong relation with the label and weak relation with each other, means they are bringing maximum number of Awardees but different from each other. Hence, $ST_{n(index)}$ for Neuroscience domain is as follows:

$ST_{n(index)} = A_{index} + K_{index}$

4.1.4 Combination of all Fields

The top five parameters from all fields, i) Authors/Paper, ii) f-index, iii) Years, iv) HI-annual, and v) K-index are extracted using correlation matrix as shown in Fig 4.4. We can see indices along x-axis and y-axis and color scheme shows relationship among these parameters. Weak relationship among two parameters means that they are bringing unique results as compare to each other. On the basis of weak relationships among each other, these five parameters are selected. Neural network back propagation approach having three hidden layers is applied on these 5 parameters to find out weights of these parameters as shown in Table 4.4. Based on the weights having the maximum impact to rank awardees, top two parameters Authors/paper and f-index are selected to formulate the index. Authors/paper and f-index had strong relation with the label and weak relation with each other, means they are bringing maximum number of Awardees but different from each

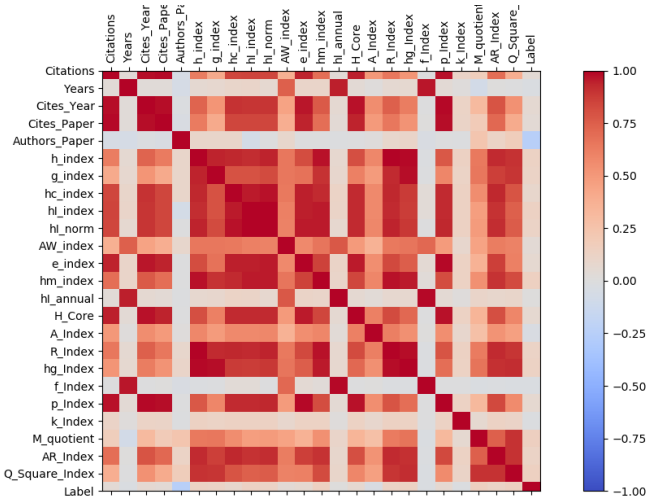

FIGURE 4.4: Correlation combined

other. Hence, ST_{index} for combination of all fields is as follows:

$ST_{index} = f_{index} + 1/(\text{Authors/paper})$

4.2 Comparison with state-of-the-art Techniques

In state-of-the-art literature, domain specific dataset is used for assessing impact of the author. In order to make an efficient index we have considered researchers from three diversified fields. To compare the results of the proposed index, we individually considered the overall result of each field. Our dataset differs from that used in the current techniques. There are imbalanced numbers of awarded and non-awardee members for the current dataset. It is not an optimal case for

TABLE 4.4: Weights All fields

Features	weights
Author/paper	0.75
f-index	0.61
years	0.29
K-index	0.17
HI-annual	0.12

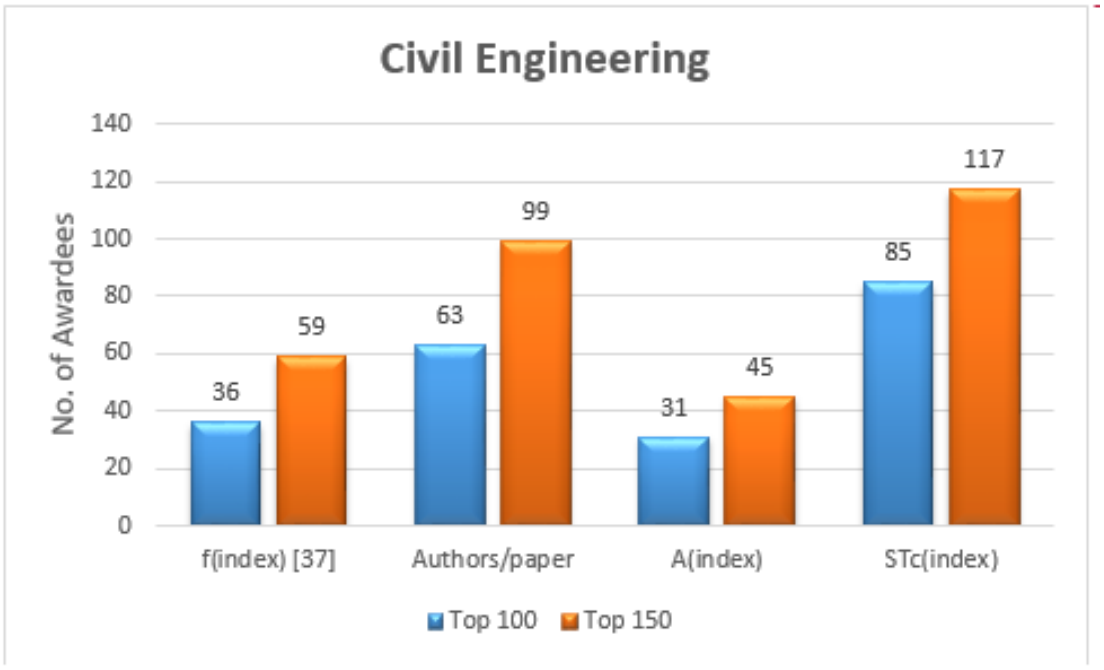

FIGURE 4.5: Results Civil Engineering

machine learning environment. So, we have taken balanced dataset. We have considered 500 researchers from each field of Civil Engineering, Mathematics and Neuroscience. Among those 500, 250 are non-awardees and 250 are awardees of prestigious scientific societies. Then we have evaluated the existing methods on this dataset to remain unbiased.

4.2.1 Civil Engineering

We have observed the occurrences of award recipients in ranked list of top 100 and top 150 researchers by the $ST_{c(index)}$, and compared it with state-of-the-art results of the existing techniques. Using the proposed index $ST_{c(index)}$, for the Civil Engineering domain, 85% are the awardees in the top 100 ranked researchers, whereas according to state-of-the-art existing technique the f-index has elevated 36% of award recipients [2].

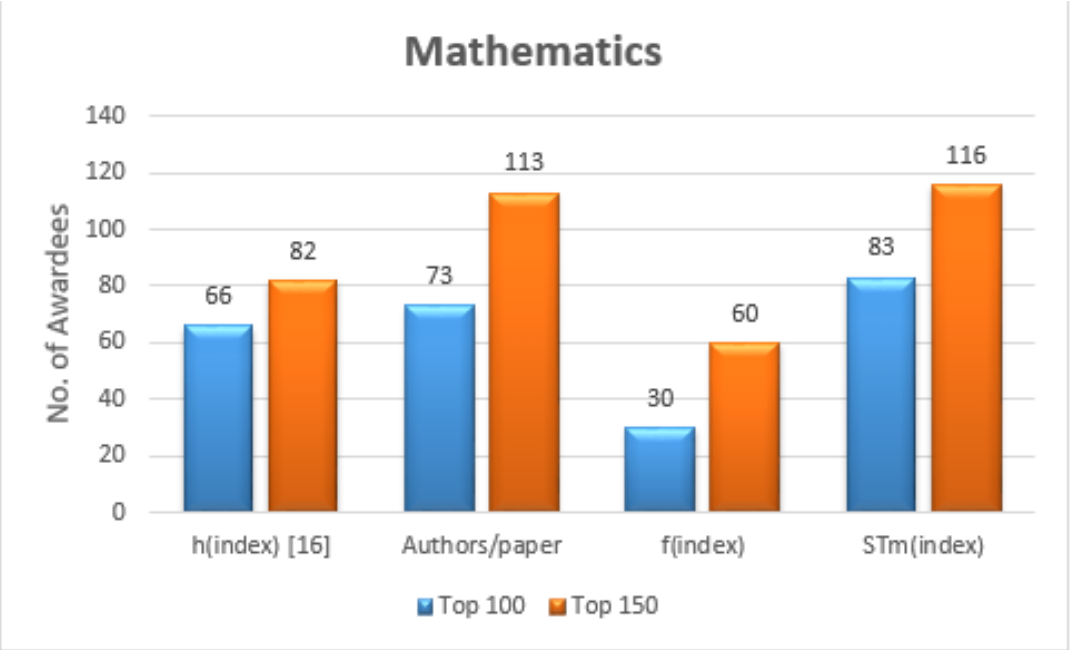

FIGURE 4.6: Results Math

4.2.2 Mathematics

We have observed the occurrences of award recipients in ranked list of top 100 and top 150 researchers by the $ST_{m(index)}$, and compared it with state-of-the-art results of the existing techniques. Using the proposed index $ST_{m(index)}$, for the Mathematics domain, 83% are the awardees in the top 100 ranked researchers, whereas according to state-of-the-art existing technique the h-index has elevated 66% of award recipients [10].

4.2.3 Neuroscience

We have observed the occurrences of award recipients in ranked list of top 100 and top 150 researchers by the $ST_{n(index)}$, and compared it with state-of-the-art results of the existing techniques. Using the proposed index $ST_{n(index)}$, for the Neuroscience domain, 85% are the awardees in the top 100 ranked researchers, whereas according to state-of-the-art existing technique the h-index has elevated 70% of award recipients [17].

4.2.4 Combination of All Fields

We have observed the occurrences of award recipients in ranked list of top 100 and top 150 researchers by the ST_{index}, and compared it with state-of-the-art results of the existing techniques. Using the proposed index ST_{index}, for the combination of all fields, 80% are the awardees in the top 100 ranked researchers, whereas according to state-of-the-art existing technique the h-index has elevated 70% of award recipients.

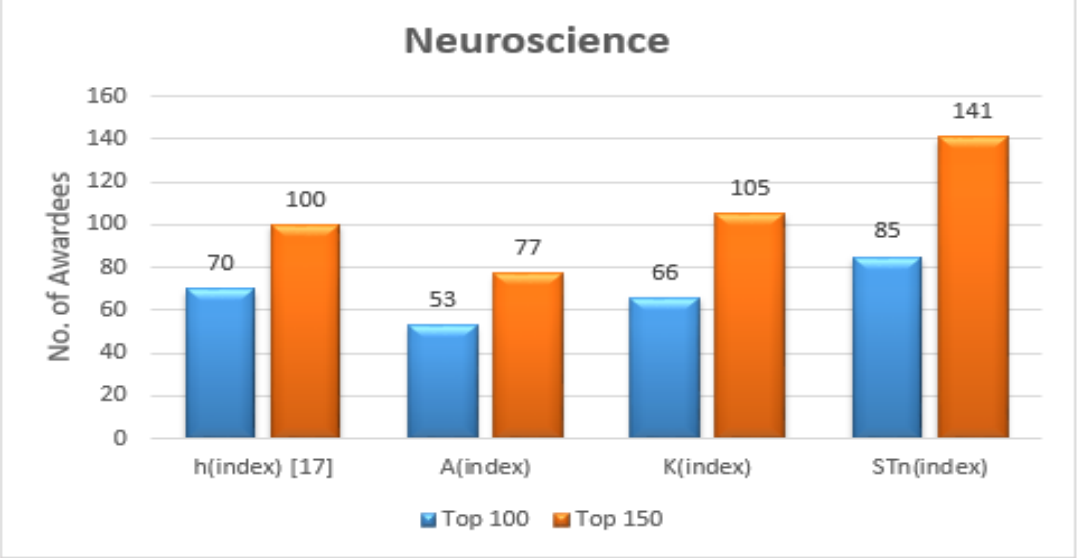

FIGURE 4.7: Results Neuroscience

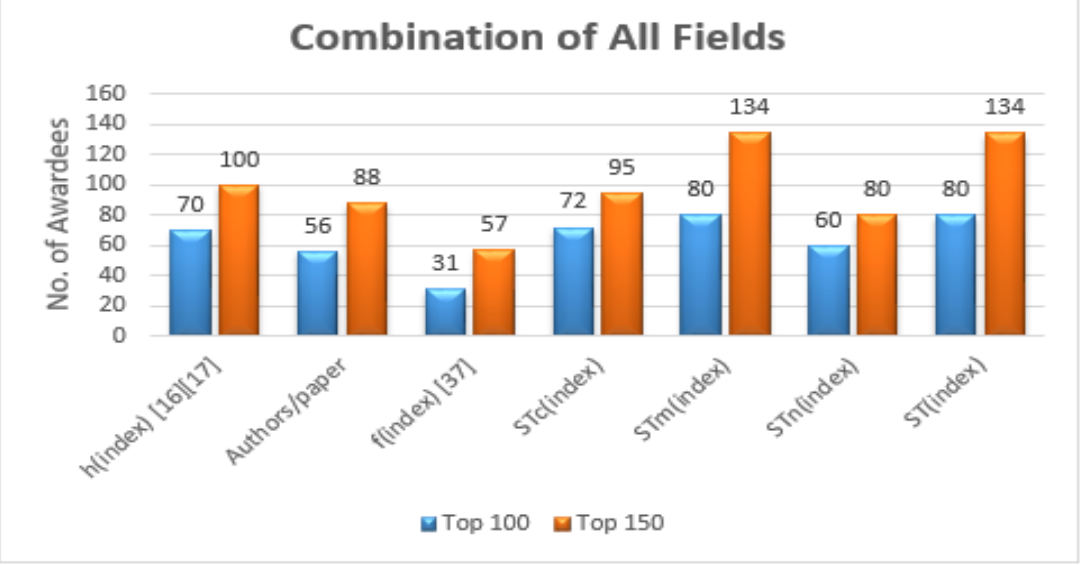

FIGURE 4.8: Results Neuroscience

Chapter 5

Conclusion and Future Work

This section will provide the conclusion of our research work and limitations for the future work.

5.1 Conclusion

The evaluation of scientist achievement of researchers has become extremely significant. There have been suggested several distinctive criteria to identify the potential of all researchers in the scientific community, including the number of publications, count of citations of the paper, h-index, A-index, R-index and other variants of h-index. The scientific community assesses the significance of researchers' work according to these parameters and classifies them as the most prolific scientist in the domain. Such criteria may also be used to reward the most impactful scientist. The students can be directed in choosing the researcher as a supervisor in order to accomplish research objectives. In academic's prospective, Universities may get help recruiting applicants with influential backgrounds in research as faculty members. It can also help the research conferences and journals to hire the member as editor or reviewer. In reference to this and many other advantages, the rank of the researchers indicates the empirical importance of the researcher to the scientific community. All of this is conceivable with the author's evaluation criteria being

taken into consideration.

From the current state-of-the-art literature analysis, it is clear that these criteria have been suggested in innovative situations and that researchers' rankings are domain-specific in literature. Furthermore, these parameters were measured on the basis of an overview of the award-winning tendency in a ranking list of scholars, without taking account of the significance of these, prior to use for a research ranking.

This research is meant to propose an index by exploring the weights of the primitive (publications, citations, years, Authors/Paper, Cites/Year, Cites/Paper, Citations in H-Core, and h-index), Citation-intensity (g-index, hg-index, A-Index, R-Index, f-index, p-index, q2- index, k-index, hm-index, and e-index, and Age-based (hc-index, hI-index, hl- norm, AW-index, hl-annual, m-quotient, and Ar-index) parameters to identify the most effective for the award recipients of Civil Engineering, Mathematics and Neuroscience.

Our first research question was to find parameters' weights which contribute effectively to rank awardees of respective fields. To answer the first research question, we have found correlation of primitive, citation-intensity and age-based parameters to analyze which author assessment parameters are bringing unique awardees at top ranking. Secondly, weights of top 5 parameters are extracted using neural network to identify the best author assessment parameters for award recipients among the all these parameters. After combining impact of the top two parameters, it is compared with the baseline papers and has achieved the accuracy of 85% to bring awardees for Civil Engineering and Neuroscience domain in the top 100 ranking. Similarly, 83% accuracy for Mathematics domain to bring awardees in the top 100 ranking. When we have combined all the researchers from Civil Engineering, Mathematics and Neuroscience domain, the proposed index has achieved 80% results.

From the above discussion it can be derived that the proposed index enabled the scientific community to determine a parameter for the author evaluation having strong relation with prominent science societies' award recipients. Moreover, ST_{index} can direct the non-awardees to be in the list of award winners, and also

helps the research groups and academic bodies in decision making about membership allocation, recruitment of editors or reviewers and faculty members.

5.2 Future Work

In addition to the primitive, citation-intensity and the age-based parameters, one major category of "Co-Author" must be evaluated on a diversified dataset. In future, we aim to evaluate the co-author-based parameters in order to conclude that the existing parameters for authors' assessment are sufficient for classifying or identifying the actual importance of researchers.

Bibliography

[1] D. James, "Completing hirsch's h-index measuring scholarly impact," *Scholardox*, vol. 74, no. 4, pp. 1014–1021, 2014.

[2] M. Raheel, S. Ayaz, and M. T. Afzal, "Evaluation of h-index, its variants and extensions based on publication age & citation intensity in civil engineering," *Scientometrics*, vol. 114, no. 3, pp. 1107–1127, 2018.

[3] L. Bornmann, R. Mutz, S. E. Hug, and H.-D. Daniel, "A multilevel meta-analysis of studies reporting correlations between the h index and 37 different h index variants," *Journal of Informetrics*, vol. 5, no. 3, pp. 346–359, 2011.

[4] R. Crowder, G. Hughes, and W. Hall, "An agent based approach to finding expertise," *Scientometrics*, vol. 104, no. 2, pp. 179–188, 2002.

[5] D. H. L. Cameron, B. Aleman-Meza, S. Decker, and I. B. Arpinar, "Semef: A taxonomy-based discovery of experts, expertise and collaboration networks," *Scientometrics*, vol. 92, no. 5, pp. 97–106, 2007.

[6] R. West and K. Stenius, "Publishing addiction science: a guide for the perplexed," *Journal of the American Socie*, vol. 71, no. 3, pp. 69–78, 2004.

[7] J. E. Hirsch, "An index to quantify an individual's scientific research output," *Proceedings of the National academy of Sciences*, vol. 102, no. 46, pp. 16 569–16 572, 2005.

[8] B. Jin, L. Liang, R. Rousseau, and L. Egghe, "The r-and ar-indices: Complementing the h-index," *Chinese science bulletin*, vol. 52, no. 6, pp. 855–863, 2007.

[9] K. R. Dienes, "Completing h," *Journal of Informetrics*, vol. 9, no. 2, pp. 385–397, 2015.

[10] S. Ayaz and M. T. Afzal, "Identification of conversion factor for completing-h index for the field of mathematics," *Scientometrics*, vol. 109, no. 3, pp. 1511–1524, 2016.

[11] L. Egghe, "Theory and practise of the g-index," *Scientometrics*, vol. 69, no. 1, pp. 131–152, 2006.

[12] B. Jin, "H-index: an evaluation indicator proposed by scientist," *Science Focus*, vol. 1, no. 1, pp. 8–9, 2006.

[13] B. Jin, L. Liang, R. Rousseau, and L. Egghe, "The r-and ar-indices: Complementing the h-index," *Chinese science bulletin*, vol. 52, no. 6, pp. 855–863, 2007.

[14] S. Alonso, F. Cabrerizo, E. Herrera-Viedma, and F. Herrera, "hg-index: A new index to characterize the scientific output of researchers based on the h-and g-indices," *Scientometrics*, vol. 82, no. 2, pp. 391–400, 2009.

[15] L. Wildgaard, J. W. Schneider, and B. Larsen, "A review of the characteristics of 108 author-level bibliometric indicators," *Scientometrics*, vol. 101, no. 1, pp. 125–158, 2014.

[16] Q.-u. Ain, H. Riaz, and M. T. Afzal, "Evaluation of h-index and its citation intensity based variants in the field of mathematics," *Scientometrics*, vol. 119, no. 1, pp. 187–211, 2019.

[17] M. Ameer and M. T. Afzal, "Evaluation of h-index and its qualitative and quantitative variants in neuroscience," *Scientometrics*, vol. 121, no. 2, pp. 653–673, 2019.

[18] L. Liang, "h-index sequence and h-index matrix: Constructions and applications," *Scientometrics*, vol. 69, no. 1, pp. 153–159, 2006.

[19] W. E. Schreiber and D. M. Giustini, "Measuring scientific impact with the h-index: a primer for pathologists," *American journal of clinical pathology*, vol. 151, no. 3, pp. 286–291, 2019.

[20] S. G. Aoun, B. R. Bendok, R. J. Rahme, R. G. Dacey Jr, and H. H. Batjer, "Standardizing the evaluation of scientific and academic performance in neurosurgery—critical review of the "h" index and its variants," *World neurosurgery*, vol. 80, no. 5, pp. 85–90, 2013.

[21] L. Smolinsky and A. Lercher, "Citation rates in mathematics: A study of variation by subdiscipline," *Scientometrics*, vol. 91, no. 3, pp. 911–924, 2012.

[22] R. Costas and M. Bordons, "The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level," *Journal of informetrics*, vol. 1, no. 3, pp. 193–203, 2007.

[23] C.-T. Zhang, "The e-index, complementing the h-index for excess citations," *PLoS One*, vol. 4, no. 5, pp. 54–29, 2009.

[24] F. J. Cabrerizo, S. Alonso, E. Herrera-Viedma, and F. Herrera, "q2-index: Quantitative and qualitative evaluation based on the number and impact of papers in the hirsch core," *Journal of Informetrics*, vol. 4, no. 1, pp. 23–28, 2010.

[25] M. Maabreh and I. M. Alsmadi, "A survey of impact and citation indices: Limitations and issues," *International Journal of Advanced Science and Technology*, vol. 40, no. 4, pp. 35–54, 2012.

[26] G. Prathap, "The 100 most prolific economists using the p-index," *Scientometrics*, vol. 84, no. 1, pp. 167–172, 2009.

[27] D. F. Thompson, E. C. Callen, and M. C. Nahata, "New indices in scholarship assessment," *American Journal of Pharmaceutical Education*, vol. 73, no. 6, pp. 11–20, 2009.

[28] A. Sidiropoulos, D. Katsaros, and Y. Manolopoulos, "Generalized hirsch h-index for disclosing latent facts in citation networks," *Scientometrics*, vol. 72, no. 2, pp. 253–280, 2007.

[29] P. D. Batista, M. G. Campiteli, and O. Kinouchi, "Is it possible to compare researchers with different scientific interests?" *Scientometrics*, vol. 68, no. 1, pp. 179–189, 2006.

[30] A.-W. Harzing, S. Alakangas, and D. Adams, "hia: An individual annual h-index to accommodate disciplinary and career length differences," *Scientometrics*, vol. 99, no. 3, pp. 811–821, 2014.

[31] C. Velmurugan and N. Radhakrishnan, "Impact of research productivity on nanotechnology in india: A scientometric profile," *International Journal of Multidisciplinary Papers,(1)*, vol. 49, no. 3, pp. 1–10, 2016.

[32] C. Wohlin, "A new index for the citation curve of researchers," *Scientometrics*, vol. 81, no. 2, pp. 521–533, 2009.

[33] A. Sidiropoulos and Y. Manolopoulos, "Generalized comparison of graph-based ranking algorithms for publications and authors," *Journal of Systems and Software*, vol. 79, no. 12, pp. 1679–1700, 2006.

[34] N. Fukuzawa, "An empirical analysis of the relationship between individual characteristics and research productivity," *Scientometrics*, vol. 99, no. 3, pp. 785–809, 2014.

[35] M. Schreiber, "Twenty hirsch index variants and other indicators giving more or less preference to highly cited papers," *Annalen der Physik*, vol. 522, no. 8, pp. 536–554, 2010.

[36] A. F. Van Raan, "Comparison of the hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups," *scientometrics*, vol. 67, no. 3, pp. 491–502, 2006.

[37] M. Raheel, S. Ayaz, and M. T. Afzal, "Evaluation of h-index, its variants and extensions based on publication age & citation intensity in civil engineering," *Scientometrics*, vol. 114, no. 3, pp. 1107–1127, 2018.

[38] A. Basu, S. K. Banshal, K. Singhal, and V. K. Singh, "Designing a composite index for research performance evaluation at the national or regional level: ranking central universities in india," *Scientometrics*, vol. 107, no. 3, pp. 1171–1193, 2016.

[39] J. Knight, "Academic impact rankings of neurosurgical units in the uk and ireland, as assessed with the h-index," *British journal of neurosurgery*, vol. 29, no. 5, pp. 637–643, 2015.

[40] M. Dunaiski, W. Visser, and J. Geldenhuys, "Evaluating paper and author ranking algorithms using impact and contribution awards," *Journal of Informetrics*, vol. 10, no. 2, pp. 392–407, 2016.

[41] H. F. Moed, J. Bar-Ilan, and G. Halevi, "A new methodology for comparing google scholar and scopus," *Journal of informetrics*, vol. 10, no. 2, pp. 533–551, 2016.

[42] R. Tol, "The h-index and its alternatives: An application to the 100 most prolific economists," *Scientometrics*, vol. 80, no. 2, pp. 317–324, 2009.

[43] L. Bornmann, R. Mutz, and H.-D. Daniel, "Are there better indices for evaluation purposes than the h index? a comparison of nine different variants of the h index using data from biomedicine," *Journal of the American society for information science and technology*, vol. 59, no. 5, pp. 830–837, 2008.

[44] F. Ye and R. Rousseau, "Probing the h-core: an investigation of the tail–core ratio for rank distributions," *Scientometrics*, vol. 84, no. 2, pp. 431–439, 2009.

[45] Q. Burrell, "Hirsch index or hirsch rate? some thoughts arising from liang's data," *Scientometrics*, vol. 73, no. 1, pp. 19–28, 2007.

[46] C.-H. Weng, T. C.-K. Huang, and R.-P. Han, "Disease prediction with different types of neural network classifiers," *Telematics and Informatics*, vol. 33, no. 2, pp. 277–292, 2016.