

Received 17 October 2023, accepted 23 November 2023, date of publication 27 November 2023,  
date of current version 30 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3336950



## RESEARCH ARTICLE

# Machine Learning Approach for Effective Ranking of Researcher Assessment Parameters

BILAL AHMED<sup>ID1</sup>, (Member, IEEE), LI WANG<sup>ID1</sup>, AHMAD SAMI AL-SHAMAYLEH<sup>ID2</sup>, MUHAMMAD TANVIR AFZAL<sup>3</sup>, (Member, IEEE), GHULAM MUSTAFA<sup>ID3</sup>, WAGDI ALRAWAGFEH<sup>ID4</sup>, AND ADNAN AKHUNZADA<sup>ID4</sup>, (Senior Member, IEEE)

<sup>1</sup>College of Information and Computer, Taiyuan University of Technology, Taiyuan 030000, China

<sup>2</sup>Department of Data Science and Artificial Intelligence, Faculty of Information Technology, Al-Ahliyya Amman University, Amman 19328, Jordan

<sup>3</sup>Department of Computing, Shifa Tameer-e-Millat University, Islamabad 44000, Pakistan

<sup>4</sup>College of Computing and IT, University of Doha for Science and Technology, Doha, Qatar

Corresponding author: Wagdi Alrawagfeh (wagdi.alrawagfeh@udst.edu.qa)

This work was supported by the Qatar National Library.

**ABSTRACT** The measurement and assessment of academic performance is now a fact of scientific life. This assessment guides the scientific community in making significant judgments such as selecting appropriate candidates for various positions, nominating individuals for scientific awards, and awarding scholarships or grants. Several research assessment parameters have been proposed by researchers to identify the most influential scholars. In the literature, researchers have employed a combination of hypothetical and fictional scenarios, as well as manual approaches, to identify the best assessment parameters. Moreover, there is no established benchmark available for assessing these parameters. The current study employs an innovative machine learning approach, the Dynamic Random Forest with Brouta Optimizer called “BorutaRanked Forest”, to prioritize the assessment metrics for researchers by calculating the importance score for each metric. Thirty different assessment metrics have been evaluated on a comprehensive dataset of researchers that contains awardees researchers and non-awardees researchers of three decades from (1990 to 2023). The main purpose of this evaluation is to determine the potential value and significance of each parameter relative to others. In addition, the position of awardees researchers is examined at different percentile ranges from Top 10% to Top 100% in the ranked lists of each parameter. During the individual evaluation of each parameter, we uncovered several intriguing patterns in the data. Our findings indicate that the normalized h-index is a particularly effective assessment parameter for the impact evaluation of researchers in the domain of mathematics. An analysis has been conducted to explore the correlation between parameters and awarding societies, examining the associations between different metrics and specific awarding societies.

**INDEX TERMS** Research evaluation, H index and variants, research assessment parameters, ranking of researchers, math subject classification.

## I. INTRODUCTION

If we do not press harder for better assessment parameters, we risk making bad funding decisions or sidelining good scientists [1]. The assessment and ranking of researchers within the scientific community has become a paramount issue [2], [3], [4]. It facilitates the scientific community to take vital decisions such as the nomination for a

The associate editor coordinating the review of this manuscript and approving it for publication was Essam A. Rashed<sup>ID</sup>.

scientific award, scholarships/grants, tenure-track positions, promotions, editor or reviewer for a journal or conference and to identify the leading professionals in a particular area [5], [6], [7]. Furthermore, it facilitates the researchers in finding a relevant research supervisor for their Ph.D. So, there's a long literary history of researchers ranking [8], [9].

A plethora of research assessment strategies have been proposed in the existing scientific literature [6], [10]. Each approach has a unique way of ranking researchers based on their contributions. The conventional approach to ranking

researchers incorporates various quantitative indicators, such as publication counts, citation counts, and the number of co-authors along with qualitative factors like peer review and expert assessment [11].

The number of publications has traditionally been the conventional measure of researchers' output [12]. Later, the research community argued that relying solely on the number of publications was an inadequate measure of a researcher's scientific impact. To overcome the limitation of publication count citations of a publication are employed as a means to gauge the influence of a researcher [13]. However, citation count also has certain limitations and shortcomings. For example, the idea of self-citation where researchers cite their own articles. In 2005, Hirsch introduced a novel parameter called the h-index [14], which combines both publications and citations into a single measure. Due to its simplicity, the scientific community widely adopted the h-index as a measure for ranking researchers and major journals. Despite its benefits, the h-index has been subject to criticism in scholarly literature due to its limitations [15], [16]. For example some of the limitations are the h-index can be bias towards researchers who predominantly publish in high-impact journals or those with a larger volume of publications in lower-quality journals. Moreover, excessive self-citations can potentially inflate the h-index, leading to an overestimation of a researcher's research impact. The h-index often overlooks highly cited articles. To address the limitations of the h-index, the g index [15] is proposed as a solution to address the issue of h index is to consider the citations of a researcher's top h core publications. While the g-index helps to overcome the limitations of the h-index, it also has other drawbacks. In particular, the g-index tends to give more weight to a small number of highly cited papers resulting in a higher g-index compared to a larger number of moderately cited papers. To mitigate the limitations of the h-index and g-index, a novel index known as the Hg index [17] was introduced to assess the impact of researchers. Several other variants are also proposed by scientific community such as A index [18], R index [19], e index [20] and F index [21] etc.

The recent studies in 2023 suggests that nearly seventy parameters have been proposed by scientists working in this field to assess and rank researchers [4], [22]. De et al. [23] stated that despite the plethora of available parameters, the scientific community does not agree regarding the most effective method for ranking researchers. The reason behind that each parameter follows its own set of criteria and no universal criteria have been established in the field for ranking researchers. Previous studies have evaluated these parameters on hypothetical or fictional scenarios and with small datasets of multiple domains [24]. In addition, since these parameters depend on and are evaluated using different datasets, it becomes challenging to understand and interpret the significance of each parameter individually [22]. However, it is important to consider the limitations of such studies and further research is required to assess

these parameters using a comprehensive and extensive dataset of a specific domain. The current state-of-the-art in ranking researchers requires a comprehensive empirical evaluation of the available parameters. Therefore, there is a significant need to measure and evaluate the h-index and its variants using a comprehensive dataset of a specific domain.

The present study evaluates thirty different parameters on a diverse dataset that includes metadata of 525 awardees researchers and an equal number of non-awardees researchers in mathematics domain. Our primary objective of evaluating these parameters in this study is to identify the most effective parameters for ranking researchers in mathematics domain.

The primary key contributions of this study are:

- We gathered an extensive dataset of awardee researchers and non-awardee researchers, spanning three eras from 1990 to 2023, in the field of mathematics. This dataset encompasses metadata about the researchers, including their publications, citations, and other relevant information.
- We implement all these parameters using the extracted metadata of researchers in the Python programming language.
- A novel machine-learning based approach called BorutaRanked Forest is employed. This approach calculates the importance score of parameters, allowing us to identify the significance of each parameter.
- We analyze the positions of award winners within the top 10%, top 20%, top 40%, top 60%, top 80%, and top 100% of the ranked lists of each parameter to understand the distribution of awardees across different percentile ranges and assess the effectiveness of each parameter in accurately identifying potential awardees.
- Furthermore, we investigate the relationship between different parameters and four prominent mathematical awarding societies to assess the criteria and preferences of these societies when selecting award recipients in the field of mathematics.
- Presenting the analysis through the presentation of outcomes and identifying the most effective metrics to evaluate researchers within mathematics domain.

The rest of the paper is organized as follows. The "Literature Review" section provides an overview of several methods used in previous studies. The "Methodology" section outlines the research approach employed to analyze these parameters. In the "Result and Discussion" section, we present the findings of our analysis while assessing the performance of each parameter. The "Conclusion" section summarizes the main findings of the study. It also identifies some limitations of the current study. Finally, the "Future Work" section discussed some potential directions for future research.

## II. LITERATURE REVIEW

In recent decades, the assessment of individual scholars and research groups has gained significant importance

[14], [23], [25]. This assessment of researchers enables the scientific community to make valuable decisions, such as selecting candidates for scientific awards, fellowships/grants, tenure-track positions, promotions, and appointments as editors or reviewers for any reputed journals or conferences. It also helps to identify prominent experts in a particular field. Furthermore, this evaluation also assists students in finding a suitable research supervisor for their doctoral studies [26], [27].

The existing scientific literature offers a wide range of research assessment strategies that have been proposed to rank researchers, including number of publications [28], number of citations [12], number of co-authors and qualitative factors such as peer review and expert assessment [29].

Publication counts have traditionally been used to measure scientific output. Although this approach works well, an author can have written very few papers, but all of them are more influential and impactful than an author who has published extensively but in low-quality journals or conferences. For example, in the field of computer science researcher Tim Berners-Lee the inventor of the World Wide Web, had a massive impact on society with his impracticable creation, despite having a relatively limited number of publications compared to other computer scientists. This parameter does not accurately measure the scientific impact of a researcher. This only represents the volume of research papers [30]. In addition, a high number of citations can be seen as a sign of influence and recognition in the scientific community. However, this metric is not necessarily a reliable indicator of the quality and longevity of a scholars work [31]. Furthermore, the practice of self-citation, where researchers cite their own work to increase the number of citations of a research paper [27].

To address the aforementioned limitations, Hirsch proposed a novel metric known as the h-index [14], which considers the number of publications and citations in a single metric. The scientific community has shown significant interest in the h-index due to its simplicity and effectiveness. Despite several advantages of the h-index the scientific community has identified several limitations [32]. The h index did not consider how many citations a researcher's most highly cited works have received. The h-index may not be suitable for novice researchers as it requires time for the publication of research papers and the subsequent increase of citations. In addition, the h-index supports senior researchers by allowing for a gradual increase in the citation count of their older research papers. To overcome the limitations of h index several new indices and variants of h index have been proposed by the scientific community. In a recent survey paper researchers stated that more than seventy parameters have been developed to rank the researchers [22]. However, with the huge number of available parameters the community does not agree upon a single parameter. The reason behind that is each parameter uses its own criteria to rank researchers [23]. In 2007, a study [33] evaluated four parameters such as the g-index, h-index, a-index, and r-index using a dataset

of 26 physicists. The study concluded that the g-index was the most reliable parameter compared to the other indices. In 2008 researchers introduced a new parameter called hm index [34] to rank researchers. Dienes in 2016 evaluated h-index, g-index, and complementary h-index for ranking of researchers in the domain of mathematics [35]. Researcher De et al., 2018, conducted an evaluation of the h-index and several of its variants in the field of civil engineering. The variants included those based on citation intensity and publication age [23]. Schreiber et al., 2019 conducted an evaluation of the h-index and some of its variants using a dataset in the field of neurosciences [36]. In 2019, Ain et al. [37] and Ghani et al. [38] conducted a systematic evaluation of citation intensity-based indices of the h-index. Their evaluation was performed on a comprehensive data set from the field of mathematics. Moreira et al., 2021 [39] conducted an evaluation of various indices on a comprehensive dataset from the field of civil engineering. The purpose of their study was to identify the most effective metrics for evaluating author performance. In a recent study, Mustafa et al., 2023a, 2023b [7], [40] evaluated Publication, Citation-based metrics and publication age based parameters on the same dataset of mathematics. Ahmed et al., 2023 [4] evaluated author count based parameters on same data set.

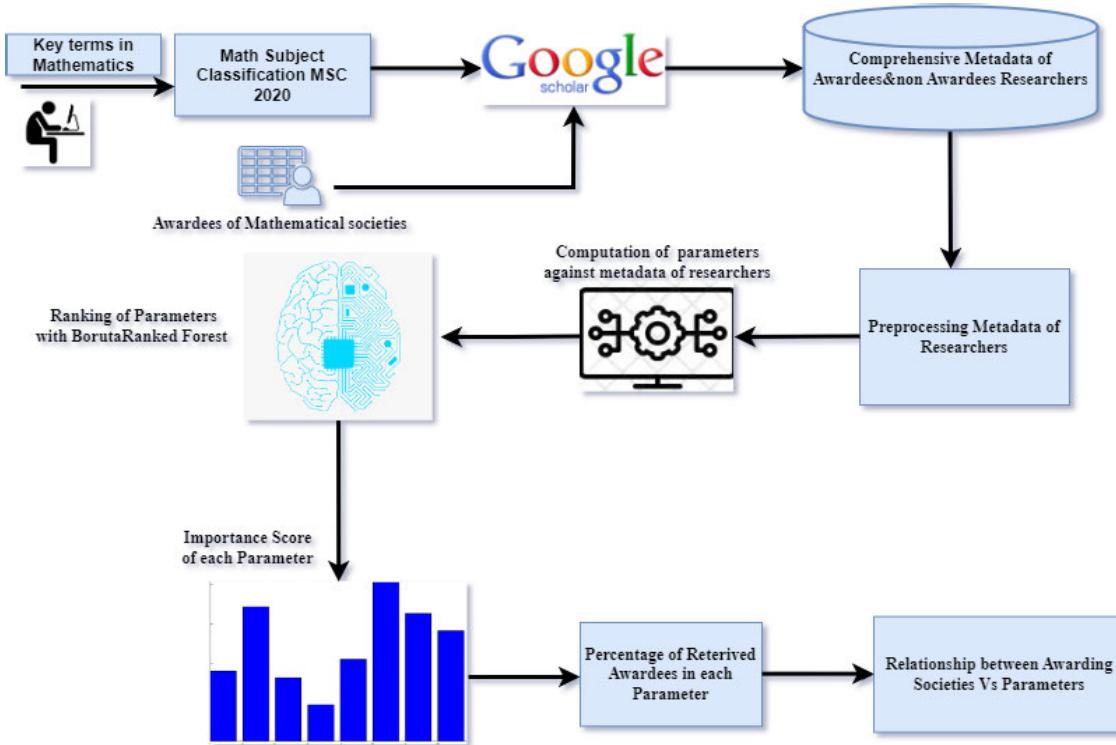
This section has reviewed several research studies in which researchers manually evaluated these parameters, without employing modern machine-learning techniques to measure the significance of these parameters. Machine learning algorithms employ various computational procedures to analyze large amounts of data to enable the identification of patterns and trends in data that indicate the potential for award winners. Therefore, we examine the role of these parameters by employing a novel modern machine learning model. By doing so, we provide the scientific community with valuable insights into the effectiveness of these parameters and their potential usage in evaluating researchers.

### III. METHODOLOGY

The scientific community has proposed a wide range of parameters to measure the scientific impact of researchers. After a comprehensive review of the current literature, we evaluate these researcher assessment parameters by employing a novel machine learning method called BorutaRanked Forest, and to analyze whether these parameters have contributed to the recognition of award winners in prestigious scientific societies. The workflow of the proposed methodology is depicted in figure 1.

#### A. COMPREHENSIVE DATASET COLLECTION IN MATHEMATICS DOMAIN

Collecting data for a particular domain typically requires the participation of domain experts. This is due to the various branches and classes within the domain, requiring specialized knowledge to ensure accurate and comprehensive data collection. Mathematics incorporates a diverse range



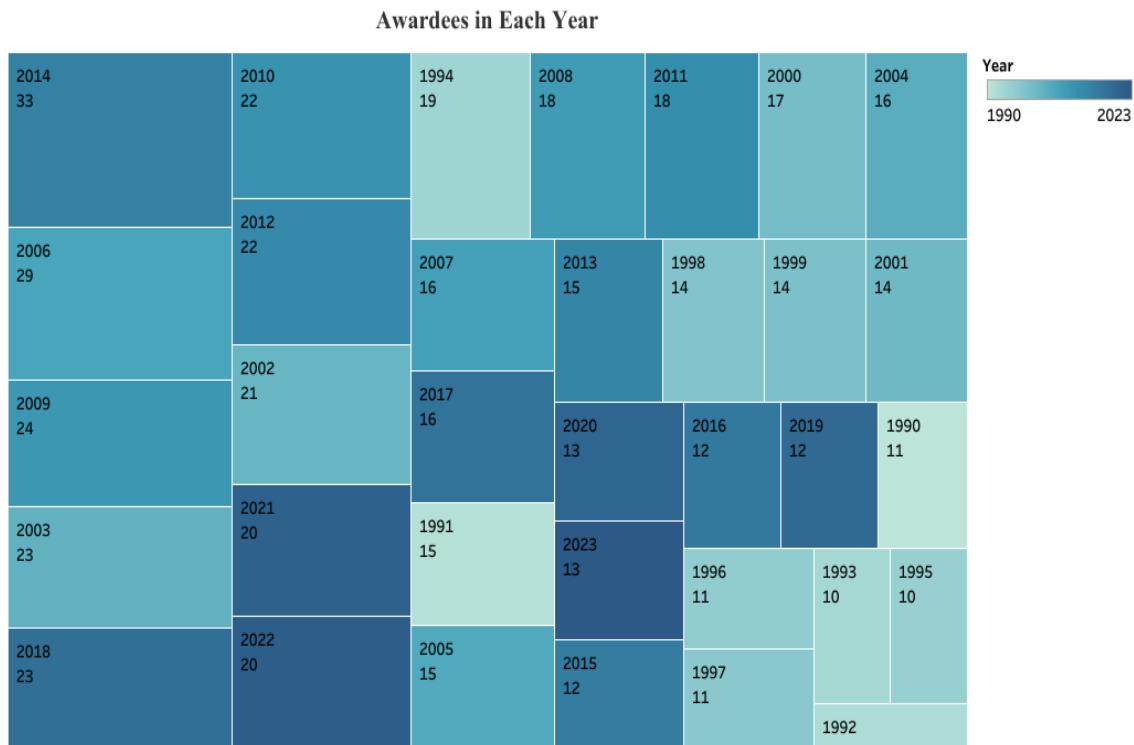
**FIGURE 1.** The block diagram of the proposed methodology.

of branches, covering numerous subfields such as algebra, geometry, calculus, probability theory, number theory, and many more. Each branch within mathematics requires specific expertise and knowledge to effectively collect data and conduct research. One source of categorizing these diverse branches of mathematics is through the Math Subject Classification (MSC) scheme [41]. The MSC scheme utilizes a hierarchical classification system to organize and classify different subfields and topics within mathematics domain. This classification system helps researchers and experts to navigate and identify several topics in the areas of mathematics. The latest version of this classification system is MSC2020. In collaboration with two renowned domain experts in mathematics, we have created a list of categories derived from the Mathematics Subject Classification (MSC). This list encompasses a comprehensive array of all the categories present in the (MSC). We manually identify several terms from the Math Subject Classification (MSC) and collect the metadata of researchers from Google Scholar. Also gathering a substantial volume of data manually and subsequently verifying its relevance to a particular domain is also a challenging task. Further we proceed to verify whether the data belongs to the domain of mathematics or not. There are multiple sources available to gather data on authors' research activities, which encompass a range of information such as publications, citations, co-author networks, and more. Some available sources include Web of Science (WOS), MathSciNet, Zbmath and Scopus. These

sources have access issues and require a subscription or membership or have limited coverage of specific disciplines or publication types. Considering these issues, we use Google Scholar database as it provides broad coverage of academic publications across diverse disciplines [42], [43]. They are accessible to researchers worldwide, enabling them to access a wide range of scholarly articles. Furthermore, the citations in Google Scholar have experienced a steady monthly growth rate of approximately 1.5%. Additionally, Google Scholar is a dynamic and continuously updated platform, regularly incorporating new data on a weekly basis [44]. This ensures that the information it offers remains current, ensuring its relevance to researchers.

### 1) BENCHMARK DATA SET

The evaluation of metrics within this ranking procedure needs a comprehensive and extensive gold standard or benchmark data set. In the specific area of research, there is no standard benchmark dataset available for the evaluation, so we have used the awardees data by scientific societies as a benchmark for our study. Many individuals receive awards in recognition of their outstanding contributions across various fields. Likewise, significant contributors and high achievers in the field of mathematics are also honored with numerous prestigious awards and distinctions. Hence, we have utilized data from researchers who have received awards from mathematical scientific societies as a benchmark for our study. The most



**FIGURE 2.** Number of Awardees in each Year.

prestigious mathematical awarding organizations include<sup>1</sup> AMS (American Mathematical Society),<sup>2</sup> LMS (London Mathematical Society),<sup>3</sup> IMU (International Mathematical Union), and<sup>4</sup> Norwegian Academy of Science and Letters, were employed as a benchmark in our analysis.

We manually search for the names of award recipients by visiting the websites of four prestigious mathematics awarding societies. Subsequently, we collect the metadata of the awardees and non-awardees spanning three decades, from 1990 to 2023, from Google Scholar. To ensure a balanced class problem and to eliminate the bias in the dataset, we have included an equal number of non-awardees in our dataset and collected non-awardee data in the same quantity for each year, corresponding to the number of awardees in those years as shown in the figure 2.

## 2) DATASET PREPROCESSING

Upon data collection, an additional stage of data refinement and validation is carried out. To achieve this, a series of meticulous steps are undertaken to ensure data quality and accuracy as shown in the figure 3.

- During the initial phase of data preprocessing, the initial step involves the removal of invalid characters, which

may include special symbols such as (\$, %, #, &, and others).

- After the removal of invalid characters, the subsequent step involves the utilization of a verification process to assess whether the papers fall within the domain of the mathematics field.
- Furthermore, author disambiguation is carried out, which includes the removal and elimination of duplicate entries and the correction of any ambiguities in authors' first or last names.

The characteristic and properties of the final dataset, following the verification of the aforementioned steps is presented in the table 1.

**TABLE 1.** Data set description for evaluation.

Researchers	Counts with Metadata
Total No of Researchers	1050
No of Awardees Researchers	525
No of Non-Awardees Researchers	525
Total Citations	15769907
Total Publications	224916

## B. COMPUTATION OF PARAMETERS

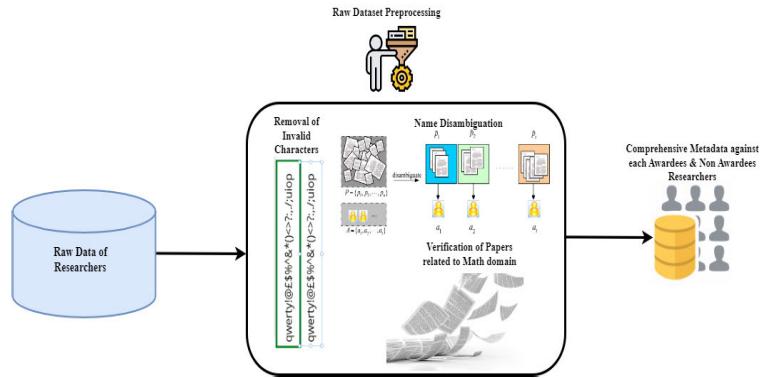
Once the metadata of a researcher has been collected, various metrics have been calculated based on this data. A Python utility is employed to calculate all of these indices. Following the computation of multiple indices on a comprehensive

<sup>1</sup><https://www.ams.org/home/page>

<sup>2</sup><https://www.lms.ac.uk/>

<sup>3</sup><https://www.mathunion.org/>

<sup>4</sup><https://dnva.no/norwegian-academy-science-and-letters>



**FIGURE 3.** Comprehensive dataset preprocessing.

dataset of researchers, we have subsequently ranked these researchers individually based on each index. We generate multiple lists for each index corresponding to each researcher, as depicted in the figure 4. The calculation method and a concise introduction to these indices are presented in this section.

#### • M-quotient

The formula of M-quotient are given below

$$M - \text{Quotient} = \frac{h - \text{index}}{y} \quad (1)$$

In above equation,  $y$  represents no of the year since the first publication and  $h$ -index represent the h-index of the author.

#### • Ar-index

The Ar-index is defined as the sum of the average number of citations per year of articles included in the h-core s. The formula of Ar-index are given below:

$$Ar - \text{index} = \sqrt{\sum_{j=1}^h \frac{Cit_j}{a_j}} \quad (2)$$

In above equation,  $Cit_j$  represent the total citation in one year,  $a_j$  represent the year and  $h$  represent the h index value.

#### • AWCR (Age Weight citation ratio)

A measure of the average number of citations for an entire body of work, adjusted for the age of each paper.

#### • Platinum H-index

The formula of Platinum H index index is given below

$$\text{platinumh} = \frac{h}{CL} * \frac{Cit_{all}}{Pub_{count}} \quad (3)$$

In above equation,  $h$  is the h-index,  $CL$  is the career length,  $Cit_{all}$  is the total citation count and  $pub_{count}$  is the publication count.

#### • hf index

The hf is a fractional counting method that maintains the original publication rank while normalizing citations. In this method, the citation count of each paper is divided

by the number of co-authors resulting in a normalized citation count. Mathematically it can be expressed as

$$\frac{Yh_f}{\phi(Yh_f)} \geq h_f \quad (4)$$

#### • gF index

This method employs fractional counting where the citation count remains unchanged, while the effective rank is determined by the publication rank. Mathematically it can be expressed as

$$gF = (\sum_{i=1}^k \frac{1}{\phi(i)})^2 \leq \sum_{i=1}^k y_i \quad (5)$$

#### • hi index

The hi-index represents the number of papers authored individually by an author that have garnered at least  $hi$  citations. Mathematically it can be expressed as

$$hi = \frac{h^2}{N_a(T)} \quad (6)$$

#### • Hm index

This is a modified version of the h-index that considers multiple co-authorship by fractionally counting papers based on the inverse of the number of co-authors. Mathematically it can be expressed as

$$r_{eff}(r) = \sum_{r'=1}^r \frac{1}{a(r')} \text{then} c(r(h_m)) \geq h_m \geq c(r(h_m) + 1) \quad (7)$$

#### • gm index

The gm-index is a modification of the g-index that takes into consideration multiple co-authorship. In this method, each article is assigned a fractional weight based on the number of co-authors it has. Mathematically it can be expressed as

$$g_m \leq C_{eff}(g_m) \text{ where } C_{eff}(r_{eff}) \\ \text{and } S_{eff}(r_{eff}) = \sum_{r=1}^{r(r_{eff})} \frac{1}{a(r)} c(r) \quad (8)$$

- Pure h index** The difference between the hi-index and the pure h-index lies in the denominator. In the hi-index, the denominator is the average number of scholars in the h-core articles, whereas in the pure h-index, the denominator is the square root of the average number of scholars in the h-core articles. Mathematically it can be expressed as

$$h_p(A) = \frac{h}{\sqrt{E(\text{author})}} \quad (9)$$

- hi norm index**

The hi-norm is a modified version of the h-index that normalizes citations based on the number of authors per paper

- A index**

The A-index of a scholar represents the average number of citations received by their h-core articles. Mathematically it is expressed as

$$A = \frac{1}{h} \sum_{p=1}^h cit_p \quad (10)$$

- Ar Index**

The Ar-index is defined as the summation of the average number of citations per year for articles included in the h-core. Mathematically it can be expressed as

$$Arindex = \sqrt{\sum_{j=1}^h \frac{cit_j}{aj}} \quad (11)$$

- g Index**

The g-index is a metric used to assess the overall impact and productivity of a researcher's published work. It is similar to the H-index, but it considers both the number of highly cited articles and the total number of citations received by the researcher.

- h dash Index**

Mathematically the h dash Index is expressed as follows

$$h' = Rh = \frac{eh}{t} \quad (12)$$

- h2 center Index**

Mathematically the h2 index is calculated as

$$h^2 \text{centreindex} = \frac{h * h}{\sum_{k=1}^m cit_k} * 100 \quad (13)$$

- h2 lower Index**

Mathematically h2 lower index is computed as

$$h^2 \text{lower} = \frac{\sum_{k=h+1}^m (cit_{k-h})}{\sum_{k=1}^m cit_k} * 100 \quad (14)$$

- hf Index**

This is a fractional counting method where the publication rank remains unchanged, and the citation count is

normalized by dividing it by the number of co-authors of each paper. It can be expressed as

$$\frac{yh_f}{\phi(yh_f)} \geq h_f \quad (15)$$

- hi Index**

The hi-index of a scholar is the ratio of the h-index and the average number of scholars in the h-core articles. Mathematically, it is defined as follows.

$$h_i = \frac{h}{Avg_a} \quad (16)$$

- Hi Index**

In the Hi-index, research papers are counted fractionally in accordance with the average number of authors of the papers contributing to the h-index. Mathematically, it is defined as follows.

$$H_i = \frac{h}{author(h)} \quad (17)$$

- hm Index**

The hm index of a researcher can be calculated as follows.

$$h_m = \max_r(r_{eff}(r) \leq c(r)) \quad (18)$$

- i10 Index**

The i10-index is another metric used to gauge the impact of a scholar and was introduced by Google Scholar in 2011. It is a simple and straightforward indexing measure obtained by counting the total number of a scholar's published papers that have received at least 10 citations.

- K dash Index**

Mathematically K dash can be expressed as.

$$k' = \frac{cit_{all} - pub_{count}}{cit_t - cit_h} \quad (19)$$

- K Index**

Citations in h-tail are not considered by h-index. To avoid this loss K-index is proposed. The rest of the publications that are not part of the H-core are significant for the K-index. Mathematically it is calculated as

$$Kindex = \frac{c/p}{c(h-tail)/c(h-core)} \quad (20)$$

- m Index**

The m-index of a scholar is the median number of citation count of the h-core articles.

- m quotient Index**

In addition to the h-index value, the M-quotient takes into account the academic age of the author. It can be calculated as

$$\frac{h - index}{y} \quad (21)$$

- **Maxord Index** The highest value is obtained by multiplying each publication index ( $i$ ) with the corresponding citation count ( $ci_i$ ) of the  $i$ th most cited paper, considering all citations.
- **Normalized h Index** The Normalized h-index is determined by dividing the h-index of a scholar by the square root of their total number of publications. This normalization allows for a fair comparison of the scholar's impact, irrespective of the varying number of publications they have. Mathematically it can be calculated as

$$\text{normalized h index} = \frac{h}{\sqrt{\text{Pub}_{\text{count}}}} \quad (22)$$

- **P Index** The p-index represents the best balance between the total number of citations (C) and the average citation rate (C/P).

$$\text{P index} = \left( \frac{C^2}{P} \right)^{1/3} \quad (23)$$

- **Pi Index** The pi-index was introduced to prioritize highly influential papers. Mathematically it can be shown as

$$\text{Pi index} = 0.01C(P\pi) \quad (24)$$

- **Platinum h Index**

The Platinum H-index includes the total number of citations, the total number of research careers and the total number of publications. It is formally defined as follows.

$$\text{Platinum h index} = \frac{H}{CL} * \frac{\text{Cit}_{\text{all}}}{\sqrt{\text{Pub}_{\text{count}}}} \quad (25)$$

- **Pure h Index**

The pure h-index is expressed as follows

$$h_p(A) = \frac{h}{\sqrt{E(\text{author})}} \quad (26)$$

- **Woginger index**

The w-index of a scholar is similar to the h-index. It is the highest value  $w$  for which their  $w$  articles have received at least  $1, 2, 3, \dots, w$  citations each. Mathematically, it is defined as follows.

$$W = \max_w (cit_{p \geq w-p+1}) \quad (27)$$

- **R index**

The R-index of scholars is calculated as the square root of the sum of citation counts for their h-core articles. Mathematically, the R-index is defined as follows.

$$R = \sqrt{\sum_{p=1}^h cit_p} \quad (28)$$

After computing all these metrics, separate ranking lists of researchers were generated for each metric as shown in the figure 4. Due to the extensive number

of parameters, only a subset of them is displayed. In the figure 4 the first column, 'Author Name,' indicates the researcher's name and the last column, 'Class,' indicates whether the researcher is an awardee ('1') or a non-awardee ('0'). All other columns represent various parameters.

### C. RANKING OF METRICS USING BORUTARANKED FOREST

To rank multiple parameters, we employ a novel BorutaRanked Forest model. In the field of scientometrics, where the primary objective frequently involves evaluating the influence of diverse scholarly contributions and comprehending intricate patterns within research data, the decision to utilize our proposed model is motivated by several factors [45]. Random Forest stands out as a robust algorithm capable of capturing non-linear relationships and interactions among variables. In the context of Scientometrics data where numerous factors can influence the importance and impact of researchers' work, Random Forest proves highly adept at modeling these complex relationships. Also the Boruta Optimizer plays a crucial role by improving the feature selection process. In scientometrics, data sets can be high-dimensional and noisy due to various metrics, publication sources, and author profiles, making feature selection crucial for identifying the most relevant contributors to research impact. The Boruta Optimizer, based on Random Forest, not only ranks feature importance but also helps in handling feature selection for large-scale data sets efficiently. The comprehensive workflow of BorutaRanked Forest is explained in detail in Algorithm 1.

This approach ranks the parameters and assigns an importance score to each parameter of researcher. The random forest classifier is designed by aggregating numerous decision trees. This procedure is executed by choosing subsets of the original data and features form each tree. Furthermore, the accuracy of each feature is subsequently assessed based on the classification results of each tree. Throughout the learning process, the least significant features are consistently eliminated in each iteration. Subsequently, the Boruta feature-selection method is applied. This method uses the random forest algorithm to enhance the feature-selection process.

Boruta combines the benefits of the dynamic random forest and performs a thorough evaluation of feature importance by comparing the importance of features with their corresponding shadow features, resulting in a comprehensive feature selection approach [46], [47]. It also considers both the absolute importance of the feature and its relative importance to shadow features, allowing for a more robust feature selection process. The algorithm Boruta calculates the Z-scores of each input predictor with respect to the attribute of the shadow. Based on the distribution of the Z-score metrics, the critical factors of the predictors are determined. To achieve an optimal feature selection approach, the Boruta

Ranked List of Awardees &amp; Non Awardees Against each Index

Author Name	Ar index	Center I.	H2 index	M Index	Normalized H Ind.	Pi Index	T Index	Woginger Index	Class
A. Corti	59.3	19.2	8.8	132.0	1.0	33.8	39.0	39.0	1.0
A. Spielman	94.2	31.6	31.5	195.0	0.6	141.5	173.0	186.0	1.0
Brooke Shipley Hrushovski	51.4	22.8	27.6	86.0	0.1	36.2	58.0	1.0	1.0
DL Powers	34.2	27.1	9.7	38.0	0.1	11.4	33.0	2.0	0.0
H Stehfest	45.9	1.8	5.6	67.0	0.1	55.0	20.0	4.0	0.0
J.B. Meleod	65.9	24.8	20.3	71.5	0.2	35.9	59.0	5.0	1.0
James Newton	186.5	21.0	30.6	359.0	0.2	936.7	318.0	37.0	1.0
Jean-christophe Yoccoz	19.7	14.5	7.7	55.5	0.5	6.1	17.0	12.0	1.0
John Huerta	140.6	9.5	11.8	559.0	0.6	313.6	132.0	118.0	1.0
Keisuke Hara Zilber	4.5	22.5	0.2	17.0	0.2	0.4	4.0	0.0	1.0
M. Owen	131.4	18.1	29.1	405.0	1.0	281.7	118.0	118.0	1.0
M.J.D. Powell	151.8	9.5	45.9	188.0	0.4	270.2	111.0	44.0	1.0
RP Agarwal	55.5	21.3	19.9	81.0	0.1	47.4	70.0	3.0	0.0
SV Shaw	28.7	24.3	10.3	50.0	0.2	13.3	36.0	4.0	0.0
U Koumba	6.2	23.5	1.4	12.0	0.3	0.5	6.0	2.0	0.0

**FIGURE 4.** Parameters with authors scores and target class.**Algorithm 1** Finding Parameters Importance Scores Using BorutaRanked Forest

**Input :** Dataset with target label  
**Output:** Importance Scores of Parameters

```

1 data ← LoadDataset;
2 X ← data.Indices;
3 y ← data.Class;
4 Xtrain, Xval, ytrain, yval ← train_test_split(X, y, test_size=0.2, random_state=42);
5 dynamic_importance ← [];
6 num_iterations ← 100;
7 boruta_selector ← Boruta(RandomForestClassifier(n_estimators=num_iterations, random_state=42));
8 for iteration in range(num_iterations) do
9   rf ← RandomForestClassifier(n_estimators=iteration + 1, random_state=42);
10  rf.fit(Xtrain, ytrain);
11  importances ← rf.feature_importances;
12  for i, importance in enumerate(importances) do
13    dynamic_importance[i].append(importance);
14  boruta_selector.fit(Xtrain, ytrain);
15 avg_importance ← np.mean(dynamic_importance);
16 indices ← np.argsort(avg_importance);
17 for i in enumerate(indices) do
18  print((data.feature_names[i]): avg_importance[i]);

```

algorithm employed a process that involved ranking the prominent Iterative Mean Filters (IMFs) and the residuals determined by Boruta factors. The Boruta algorithm for computing importance score is described in the following steps.

- Step 1:

Create a randomly ordered duplicate variable,  $\chi'_a$  for a specific input vector e.g (g index, A index, ... etc),  $\chi_b$  in order to introduce randomness and mitigate correlations between duplicate predictors and targets, ( $y_t$ ) e.g (Class Awardee “1” Nonawardee “0”) for a set of discrete inputs,  $x_t \in R^n$ , H and a target variable,  $y_t \in R$  with multiple inputs (n) and  $t = 1, 2, 3 \dots H$ . Using the random forest algorithm, the target  $y_t$  will be predicted with the duplicated inputs  $\chi'_a$  and actual input ( $x_t$ ).

- Step 2:

The variance significance measures i.e., Mean decrease accuracy (MDA) for every input ( $x_t$ ) e.g (original feature) and corresponding shadow input e.g (shadow feature)  $\chi'_a$ . The complete tree size used in this analysis is 500 i.e  $m_{tree} = 500$  as shown in equation.

$$MDA = \frac{1}{m_{tree}} \sum_{m=1}^{m_{tree}} \frac{\sum_{t \in OOB} I(y_t=f(x_t)) - \sum_{t \in OOB} I(y_t=f(x_t'))}{|OOB|}$$

In the above equation  $I(\bullet)$  is an indicator function “OOB” is denoted as Out-of-Bag is the predicted error of each training sample based on bootstrapping aggregation whereas ( $y_t = f(x_t)$ ) are prediction values

before mean decrease accuracy and ( $y_t = f(x_t^n)$ ) is prediction values after mean decrease accuracy.

- Step 3:  
The Z score is computed as follows

$$Z - Score = \frac{MDA}{SD}$$

In the above equation, SD represents the standard deviation of the accuracy losses, and the maximum Z-score is calculated for the shadow features. The Z-scores of the predictors are then compared with their corresponding duplicates and analysed using the variable importance distribution.

- Step 4:  
New duplicate inputs are generated, and the algorithm terminates either when all input parameters are confirmed or when the iteration threshold is reached.

The Boruta algorithm is a robust feature-selection method that proves especially valuable for datasets characterized by a high number of irrelevant or redundant features. Furthermore, it exhibits resistance to overfitting and possesses the capability to handle both continuous and categorical data. This algorithm has demonstrated its effectiveness in a range of applications, including but not limited to bioinformatics, genetics, and scientometrics data [48], [49].

The comprehensive workflow to rank these parameters is shown in figure 5.

#### IV. RESULTS AND DISCUSSIONS

Feature ranking plays a crucial role in machine learning as it aids in identifying the most influential features for prediction, enhancing model interpretability, and reducing dimensionality. In our case we calculate the importance of each researcher assessment metric like m index, k index, g index e.t.c. Following the application of the aforementioned feature selection method, we proceed to compute the importance score for each metric. Thirty different variants of the h index have been evaluated in this study. The figure 6 clearly demonstrates that the normalized h index has outperformed all other parameters, achieving the highest rank with an accuracy of 60%. In comparison, the gm index obtained a rank with an accuracy of 55%, while the h2 index achieved a rank with 53% accuracy. The poor performance is shown in the case of the Ar index.

Once the importance score for each metric has been computed, we conduct an analysis to determine the percentage of awardees captured by each metric within the ranked list across various percentile ranges. We have examined the percentage of awardees within different percentile ranges of the ranked lists, specifically at the top 10%, top 20%, top 40%, top 60%, top 80%, and top 100%.

It has been commonly believed that award recipients typically possess a strong research background, characterized

by a high number of publications and citations. As a result, there has been the probability that all awardees would consistently rank within the top 10% of authors when sorted by these indices, given their established strong research background. However, recent analyses have shown that this assumption is not always valid, as there have been cases where certain award recipients do not meet the expectation of ranking within the top 10% based on these indices.

From the figure 7 we can analyze that within the top 10% of the ranked list, the A index, gF index, and normalized h index have successfully identified 80 percent of the awardees. On the other hand, the Hi index, pi index, and R index demonstrate an average performance, capturing 50 percent of the awardees. However, the Ar index, HI index, and Maxpord index exhibit poor performance in identifying awardees.

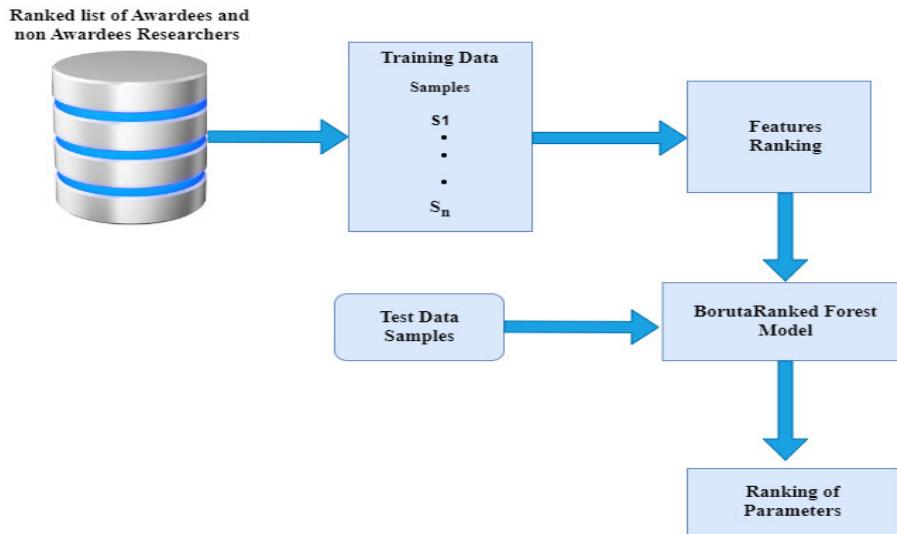
From the figure 8 we can analyze that within the top 20% of the ranked list, the normalized h index, gm index, and A index have successfully identified above 75 percent of the awardees. On the other hand, the pure h index, R index, and hf index demonstrate an average performance, capturing 50 percent of the awardees. However, the i10 index and g index exhibit poor performance in identifying awardees.

From the figure 9 we can analyze that within the top 40% of the ranked list, the normalized h index, h2 lower index, and woginger index have successfully identified above 80 percent of the awardees. On the other hand, the k index, m index, and hi index demonstrate an average performance, capturing above 50 percent of the awardees. However, the i10 index and g index exhibit poor performance in identifying awardees.

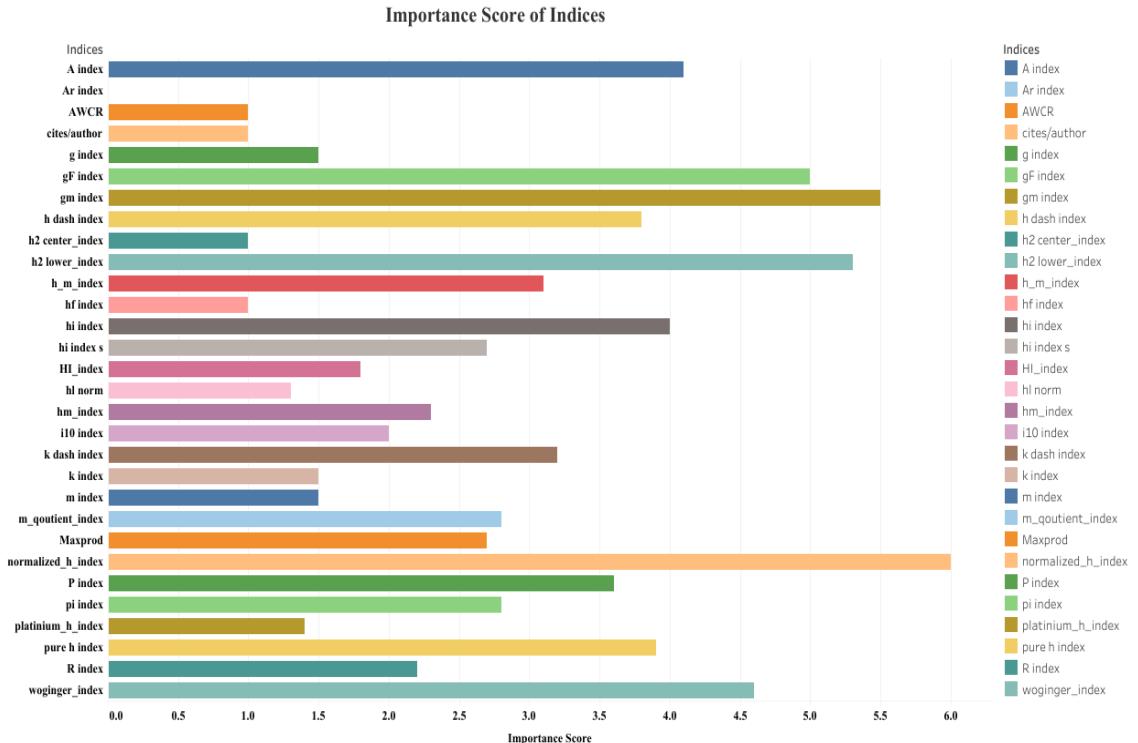
From the figure 10 we can analyze that within the top 60% of the ranked list, the normalized h index, gm index, and A index have successfully identified above 80 percent of the awardees. On the other hand, the h dash index, k index, and p index demonstrate an average performance, capturing above 50 percent of the awardees. However, the i10 index exhibits poor performance in identifying awardees.

From the figure 11 we can analyze that within the top 80% of the ranked list, the normalized h index, woginger index, and gF index have successfully identified above 80 percent of the awardees. On the other hand, the platinum h index, m quotient index, and pure h index demonstrate an average performance, capturing above 50 percent of the awardees. However, the i10 index and g index exhibit poor performance in identifying awardees.

From the figure 12 we can analyze that within the top 100% of the ranked list, the normalized h index, gm index, and woginger index have successfully identified above 80 percent of the awardees. On the other hand, the k dash index, k index, and hi index demonstrate an average performance, capturing 50 percent of the awardees. However, the i10 index and g index exhibit poor performance in identifying awardees.



**FIGURE 5.** Workflow diagram Ranking of parameters for importance score calculation.

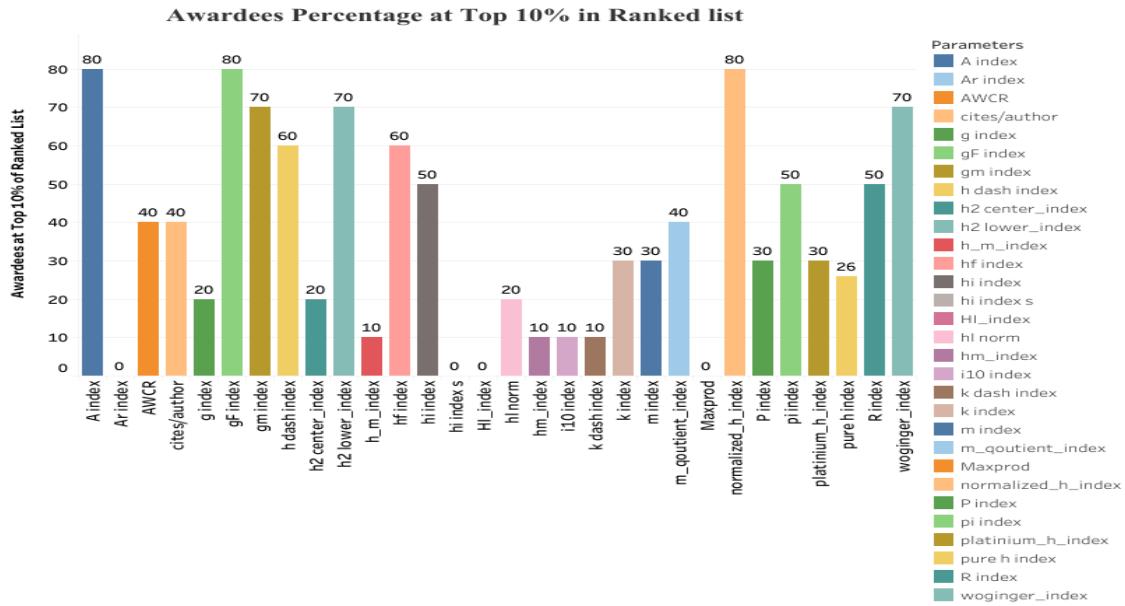


**FIGURE 6.** Importance Score of parameters.

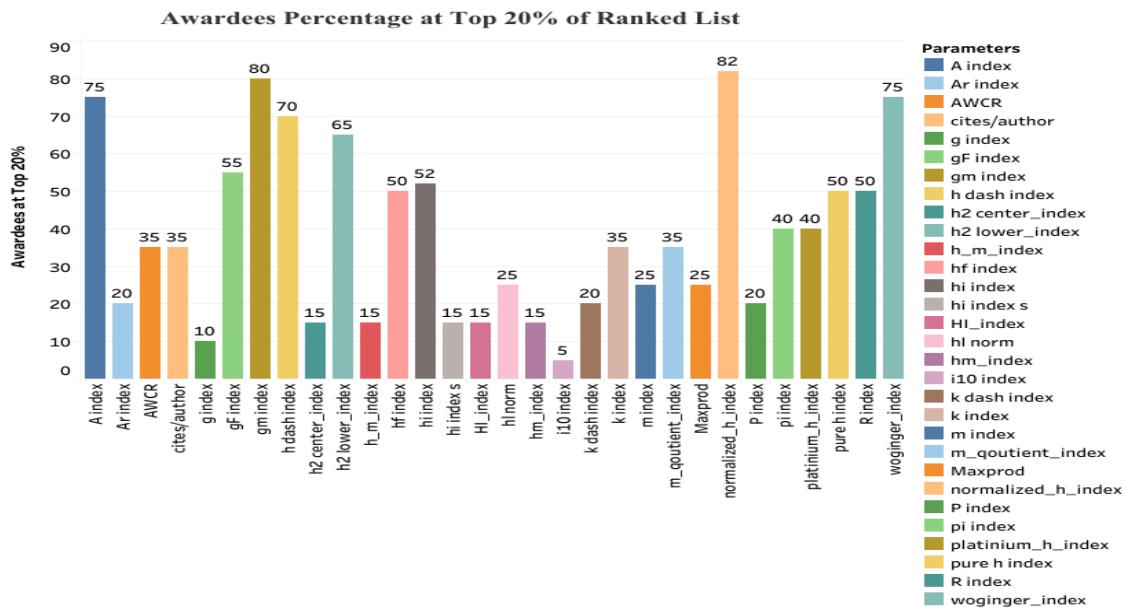
## 1) SUMMARY OF RESULTS FOR VARIOUS POTENTIAL INDICES

In this section we provide a brief overview of potential metrics in retrieving the awardees. We recorded the positions of each awardees on the ranking lists and determined the count of awardees within the top 10%, to top 100% of each metric list. For example, let's take a list of 100 researchers from a particular dataset, with 20 of them being recipients

of awards. The authors' names are organized in descending order, and subsequently, 10% of the dataset is selected. This means that this 10% segment of the data contains details about the top 10 researchers. Within the top 10 percent of ranked list “A index, gF index, and normalized h index” perform notably well among all other metrics in retrieving the 80 percent of award winners. However, when considering the entire top 100 percent of the list, the normalized h index demonstrates



**FIGURE 7.** Percentage of retrieved awardees against Indices.



**FIGURE 8.** Percentage of retrieved awardees against Indices.

the best performance, capturing over ninety percent of award winners.

## A. ASSOCIATION BETWEEN AWARDING SOCIETIES AND METRICS

In this section, we present the results of several metrics in relation to four mathematics awarding societies. We examine the dependency of each metric on different mathematics awarding societies. For this purpose, we investigated the frequency of award winners within different percentile

ranges, specifically the top 10%, top 50%, and top 100%. Based on our analysis of the dependency of awarding societies on these parameters, we have made the following observations.

## 1) AMERICAN MATHEMATICAL SOCIETY (AMS)

In the top 10% of the ranked list, the m quotient index and A index successfully identified 100 percent of the awardees. Additionally, the Maxpord index, h2 center index, and Ar index captured over 40 percent of the awardees.

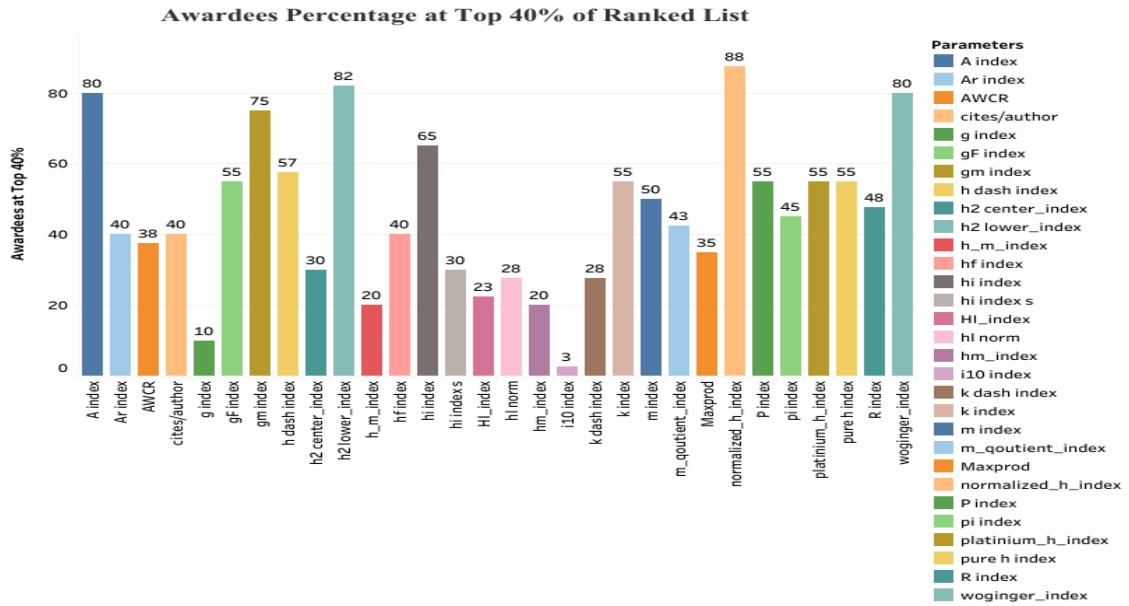


FIGURE 9. Percentage of retrieved awardees against Indices.

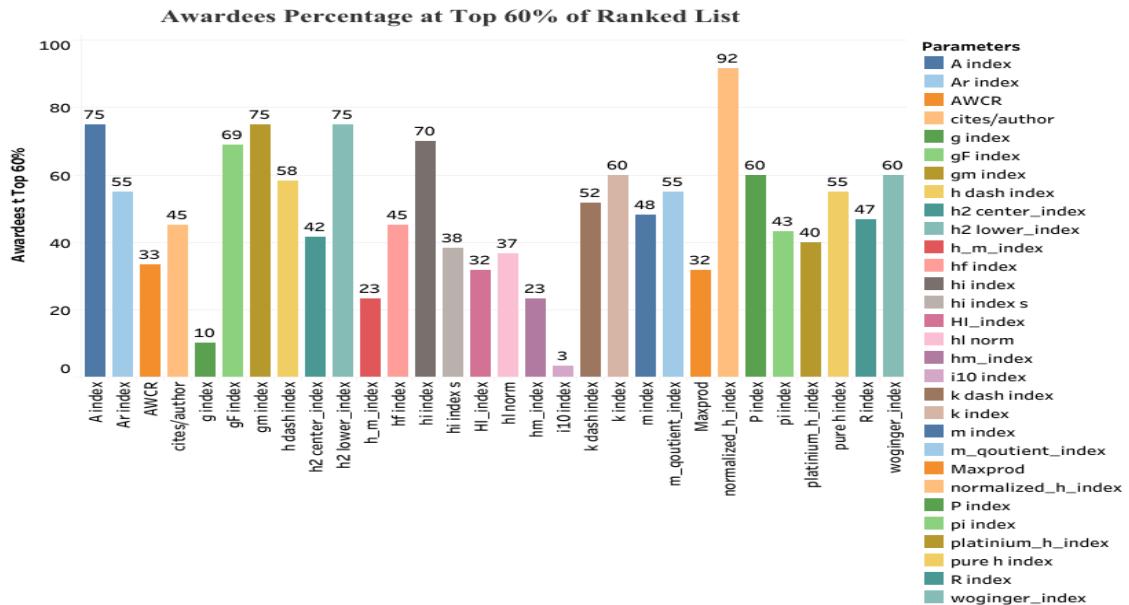


FIGURE 10. Percentage of retrieved awardees against Indices.

However, poor performance was observed for indices such as hm index, g index, and platinum h index, as depicted in the figure 13.

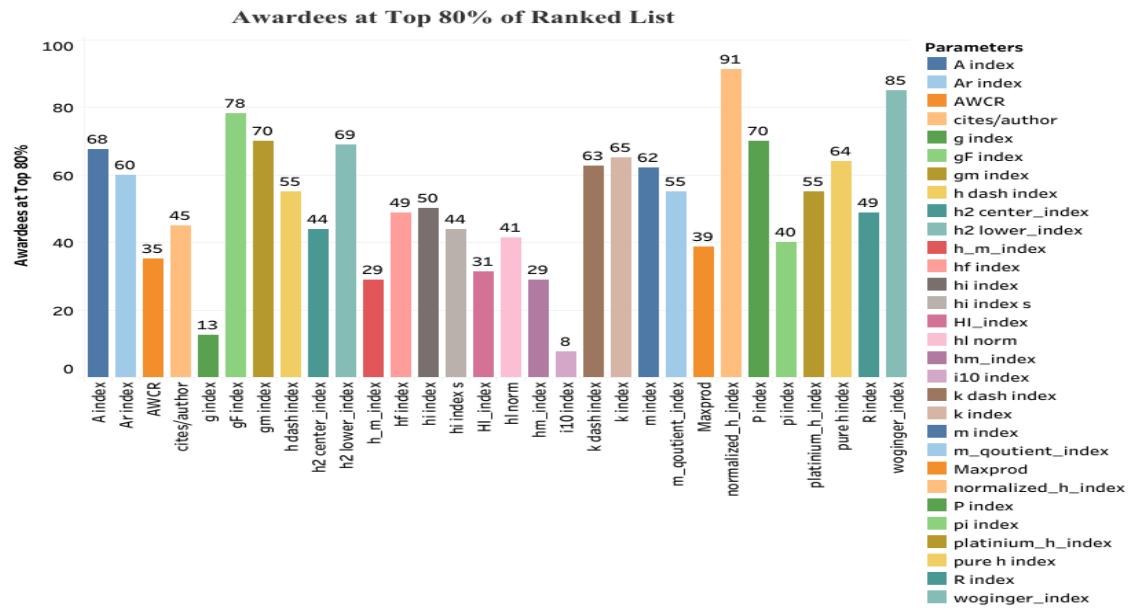
In the top 50% of the ranked list, the hi norm index and m index successfully identified 80 percent of the awardees. Additionally, the g index, h2 lower index, and hi index captured over 50 percent of the awardees. However, poor performance was observed for indices such as hm index, gF index, and R index, as depicted in the figure 14.

In the top 100% of the ranked list, the R index and m index successfully identified 70 percent of the awardees.

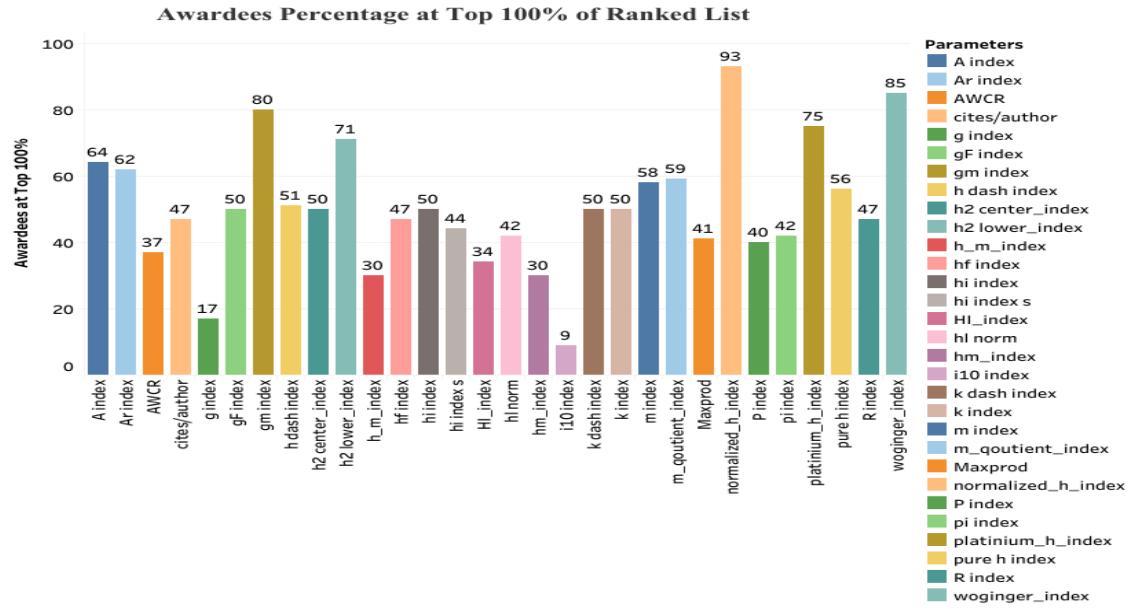
Additionally, the hi index and hi norm index, captured over 50 percent of the awardees. However, poor performance was observed for indices such as A index, h dash index, as depicted in the figure 15.

## 2) INTERNATIONAL MATHEMATICAL UNION (IMU)

In the top 10% of the ranked list, the hf index and R index and hi norm index successfully identified 100 percent of the awardees. Additionally, the m index, i10 index captured over 50 percent of the awardees. However, poor performance was



**FIGURE 11.** Percentage of retrieved awardees against Indices.



**FIGURE 12.** Percentage of retrieved awardees against Indices.

observed for indices such as g index, normalized index, and gF index, etc. as depicted in the figure 13.

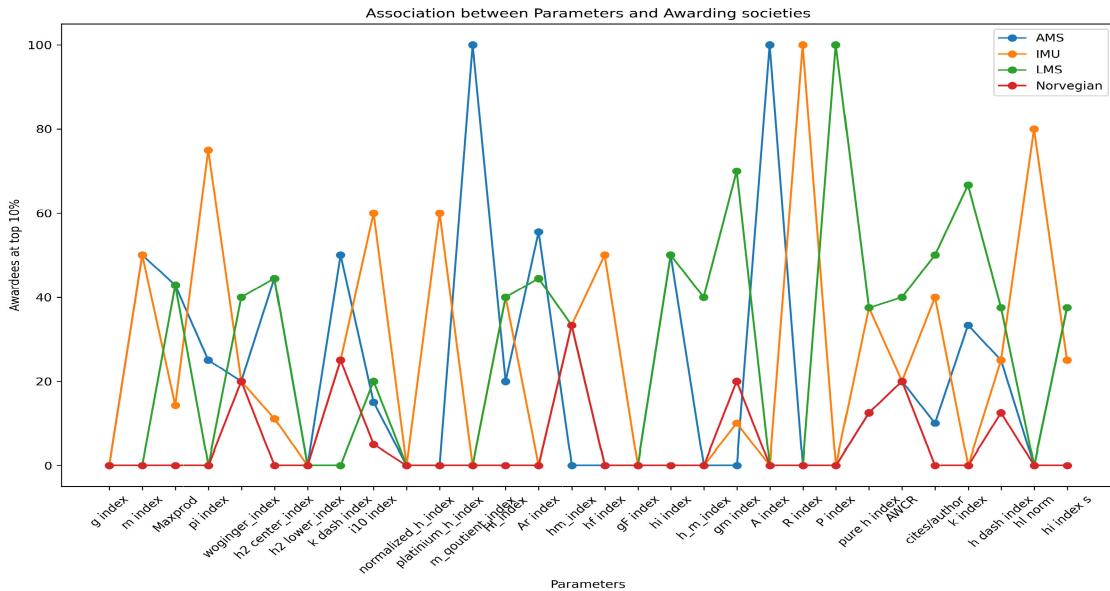
In the top 50% of the ranked list, the R index and gF index successfully identified above 80 percent of the awardees. Additionally, the hi index s, hi norm index captured over 20 percent of the awardees. However, poor performance was observed for indices such as hm index, as depicted in the figure 14.

In the top 100% of the ranked list, the normalized h index successfully identified 30 percent of the awardees. Additionally, the hi index, woginger index captured over

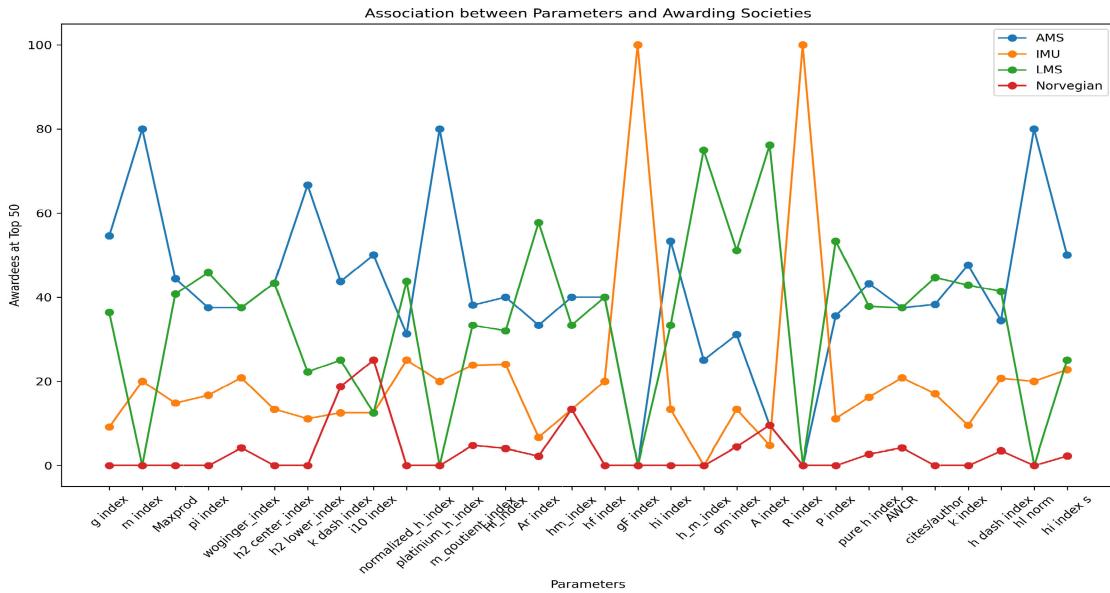
17 percent of the awardees. However, poor performance was observed for g index as depicted in the figure 15.

### 3) LONDON MATHEMATICAL SOCIETY (LMS)

In the top 10% of the ranked list, the hm index and P index successfully identified 100 percent of the awardees. Additionally, the maxprod index, and h2 center index captured over 40 percent of the awardees. However, poor performance was observed for indices such as hi norm index, R index, and A index, etc. as depicted in the figure 13.



**FIGURE 13.** Association between metrics and awarding societies.



**FIGURE 14.** Association between metrics and awarding societies.

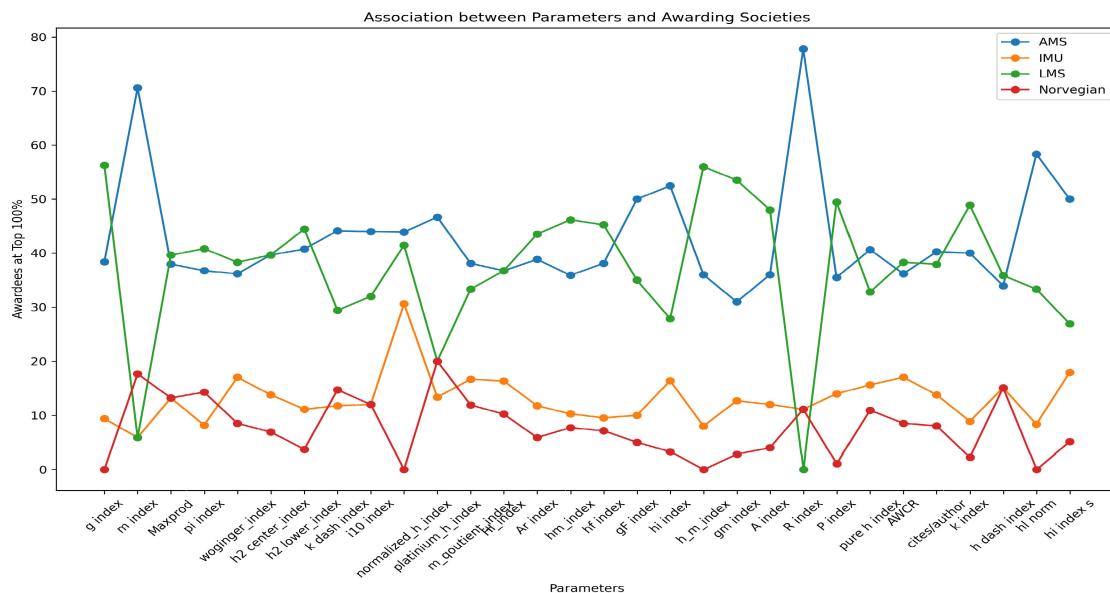
In the top 50% of the ranked list, the hm index and A index successfully identified 70 percent of the awardees. Additionally, the maxprod index, Pi index and Ar index captured over 50 percent of the awardees. However, poor performance was observed for indices such as platinum h index, and gF index, etc. as depicted in the figure 14.

In the top 100% of the ranked list, the gm index and g index successfully identified 50 percent of the awardees. Additionally, the Pi index, and h2 lower index captured over 40 percent of the awardees. However, poor performance

was observed for index such as R index as depicted in the figure 15.

#### 4) NORWEGIAN ACADEMY OF SCIENCE AND LETTERS

In the top 10% of the ranked list, the hm index successfully identified 33 percent of the awardees. Additionally, the gm index and pure h index captured over 20 percent of the awardees. However, poor performance was observed for indices such as hi norm index, k index, and A index, etc. as depicted in the figure 13.



**FIGURE 15.** Association between metrics and awarding societies.

In the top 50% of the ranked list, the k dash index successfully identified 18 percent of the awardees. Additionally, the gm index and AWCR index captured over 4 percent of the awardees. However, poor performance was observed for indices such as hi norm index, cites/author, and R index, etc. as depicted in the figure 14.

In the top 100% of the ranked list, the pi index and k dash index successfully identified 14 percent of the awardees. Additionally, the woginger index and h2 center index captured below 10 percent of the awardees. However, poor performance was observed for indices such as g index, normalized h, and hm index, etc. as depicted in the figure 15.

## 5) SUMMARY OF RESULTS FOR VARIOUS SOCIETIES VS INDICES

In this section, we offer a brief overview of several indices in the context of different mathematical awarding societies. The R index and m index are suitable metrics for the American Mathematical Society (AMS) when it comes to identifying influential researchers in the field of mathematics. For the International Mathematical Union (IMU) society, we have determined that the normalized h index is well-suited for recognizing influential researchers for this awarding society. Furthermore, when considering the London Mathematical Society (LMS), both the gm index and g index prove highly suitable metrics. In the case of the Norwegian Academy of Science and Letters (NASL) society, the pi index and k dash index are deemed satisfactory choices.

## V. CONCLUSION

Evaluating the scientific influence of a scholar holds great importance due to the numerous benefits. Several research

assessment parameters have been proposed in the literature for the acknowledgement of most influential researchers. These assessment parameters can be utilized to rank or recruit researchers in a specific domain. The current state of the art literature suggests that these assessment parameters are often developed and evaluated using hypothetical or fictional cases as a basis for analysis and evaluation. Furthermore, there is a deficiency of standardized benchmark datasets for assessing these indices.

We have evaluated these parameters on a comprehensive data set of researchers of mathematics domain. We ranked each assessment parameters using a novel BorutaRanked Forest. Subsequently, we will examine the percentage of retrieved awardees within each metric across different percentile ranges, specifically at the Top 10% to Top 100% of the ranked list. While evaluation we found that normalized h index performs well in the domain of mathematics. However, the i10 index exhibits lower performance compared to other indices. Furthermore, an analysis have been conducted to find the association between awarding societies and various metrics. This analysis, helps us to discern which awarding societies depend on specific indices when choosing influential researchers for recognition.

### A. LIMITATIONS OF STUDY

Despite the multitude of metrics and parameters proposed in the field to quantify the scientific impact in science, there is still a lack of universally accepted criteria or metrics. While the results indicate that the normalized h-index is a suitable metric for the mathematics domain, it may not necessarily perform as effectively in other scientific domains. Our study acknowledges the diversity among scientific fields and their specific evaluation criteria. The intention behind our research

was not to propose a universal, one-size-fits-all evaluation method. Instead, we aimed to explore the effectiveness of certain metrics in a specific context—evaluating awardees. In this context, we focused on a group of Mathematicians as a case study to assess the suitability of these metrics for this particular group. We recognize that the findings from our study are not intended to be extrapolated as a universal standard for evaluating all sciences. Therefore, the scope of our study is limited to the evaluation of awardees within the field of mathematics.

## VI. FUTURE WORK

In addition to these metrics, researchers and ranking communities have developed several other parameters. For future studies, our goal is to evaluate more parameters using comprehensive datasets from other domains such as Computer Science, Medical Sciences, and Engineering to find more potential metrics for that domain.

## ACKNOWLEDGMENT

The authors would like to thank Al-Ahliyya Amman University and the University of Engineering and Technology Peshawar, Pakistan, for their invaluable assistance and support, which greatly facilitated the execution of this study. Their contributions were instrumental in the successful completion of this research work.

## REFERENCES

- [1] J. Lane, "Let's make science metrics more scientific," *Nature*, vol. 464, no. 7288, pp. 488–489, Mar. 2010.
- [2] X. Jiang, X. Sun, and H. Zhuge, "Graph-based algorithms for ranking researchers: Not all swans are white!" *Scientometrics*, vol. 96, no. 3, pp. 743–759, Sep. 2013.
- [3] G. Mustafa, A. Rauf, B. Ahmed, M. T. Afzal, A. Akhunzada, and S. Z. Alharthi, "Comprehensive evaluation of publication and citation metrics for quantifying scholarly influence," *IEEE Access*, vol. 11, pp. 65759–65774, 2023.
- [4] B. Ahmed, L. Wang, G. Mustafa, M. T. Afzal, and A. Akhunzada, "Evaluating the effectiveness of author-count based metrics in measuring scientific contributions," *IEEE Access*, vol. 11, pp. 101710–101726, 2023.
- [5] 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, Mar. 2018.
- [6] B. Ahmed and W. Li, "Discretization based framework to improve the recommendation quality," *Int. Arab J. Inf. Technol.*, vol. 18, no. 3, pp. 365–371, May 2021.
- [7] G. Mustafa, A. Rauf, A. S. Al-Shamayleh, B. Ahmed, W. Alrawagfeh, M. T. Afzal, and A. Akhunzada, "Exploring the significance of publication-age-based parameters for evaluating researcher impact," *IEEE Access*, vol. 11, pp. 86597–86610, 2023.
- [8] B. Ahmed, L. Wang, W. Hussain, M. A. Qadoos, Z. Tingyi, M. Amjad, S. Badar-ud-Duja, A. Hussain, and M. Raheel, "Optimal rating prediction in recommender systems," in *Proc. 6th Int. Conf. Data Sci. (ICDS) Conf.*, Ningbo, China. Cham, Switzerland: Springer, May 2020, pp. 331–339.
- [9] A. Hussain, Y. Qiang, M. Abdul, K. Wu, Z. Zhao, and B. Ahmed, "Region-wise ranking for one-day international (ODI) cricket teams," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 10, pp. 39–46, 2019.
- [10] A. Harzing and R. van der Wal, "A Google Scholar h-index for journals: An alternative metric to measure journal impact in economics and business," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 60, no. 1, pp. 41–46, Jan. 2009.
- [11] E. Garfield, "Citation analysis as a tool in journal evaluation: Journals can be ranked by frequency and impact of citations for science policy studies," *Science*, vol. 178, no. 4060, pp. 471–479, Nov. 1972.
- [12] S. P. Nerur, A. A. Rasheed, and V. Natarajan, "The intellectual structure of the strategic management field: An author co-citation analysis," *Strategic Manage. J.*, vol. 29, no. 3, pp. 319–336, Mar. 2008.
- [13] H. Lima, T. H. P. Silva, M. M. Moro, R. L. T. Santos, W. Meira, and A. H. F. Laender, "Aggregating productivity indices for ranking researchers across multiple areas," in *Proc. 13th ACM/IEEE-CS Joint Conf. Digit. Libraries*, Jul. 2013, pp. 97–106.
- [14] K. Balog, L. Azzopardi, and M. de Rijke, "Formal models for expert finding in enterprise corpora," in *Proc. 29th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Aug. 2006, pp. 43–50.
- [15] J. Ding, C. Liu, and G. A. Kandonga, "Exploring the limitations of the h-index and h-type indexes in measuring the research performance of authors," *Scientometrics*, vol. 122, no. 3, pp. 1303–1322, Mar. 2020.
- [16] J. E. Hirsch, "An index to quantify an individual's scientific research output," *Proc. Natl. Acad. Sci. USA*, vol. 102, no. 46, pp. 16569–16572, Nov. 2005.
- [17] F.-F. Cheng, Y.-W. Huang, D.-C. Tsaih, and C.-S. Wu, "Trend analysis of co-authorship network in library hi tech," *Library Hi Tech*, vol. 37, no. 1, pp. 43–56, Mar. 2019.
- [18] L. Egghe, "Theory and practise of the g-index," *Scientometrics*, vol. 69, no. 1, pp. 131–152, Oct. 2006.
- [19] M. M. U. Rathore, M. J. J. Gul, A. Paul, A. A. Khan, R. W. Ahmad, J. J. P. C. Rodrigues, and S. Bakiras, "Multilevel graph-based decision making in big scholarly data: An approach to identify expert reviewer, finding quality impact factor, ranking journals and researchers," *IEEE Trans. Emerg. Topics Comput.*, vol. 9, no. 1, pp. 280–292, Jan. 2021.
- [20] Q. L. Burrell, "On the h-index, the size of the Hirsch core and Jin's A-index," *J. Informetrics*, vol. 1, no. 2, pp. 170–177, Apr. 2007.
- [21] N. R. Khan, C. J. Thompson, D. R. Taylor, K. S. Gabrick, A. F. Choudhri, F. R. Boop, and P. Klimo, "Part II: Should the h-index be modified? An analysis of the m-quotient, contemporary h-index, authorship value, and impact factor," *World Neurosurg.*, vol. 80, no. 6, pp. 766–774, Dec. 2013.
- [22] A. Bihari, S. Tripathi, and A. Deepak, "A review on h-index and its alternative indices," *J. Inf. Sci.*, vol. 49, no. 3, pp. 624–665, Jun. 2023.
- [23] N. De, S. M. A. Nayem, and A. Pal, "F-index of some graph operations," *Discrete Math., Algorithms Appl.*, vol. 8, no. 2, Jun. 2016, Art. no. 1650025.
- [24] F. J. Cabrerizo, S. Alonso, E. Herrera-Viedma, and F. Herrera, " $q^2$ -index: Quantitative and qualitative evaluation based on the number and impact of papers in the Hirsch core," *J. Informetrics*, vol. 4, no. 1, pp. 23–28, Jan. 2010.
- [25] G. Mustafa, M. Usman, L. Yu, M. T. Afzal, M. Sulaiman, and A. Shahid, "Multi-label classification of research articles using Word2Vec and identification of similarity threshold," *Sci. Rep.*, vol. 11, no. 1, Nov. 2021, Art. no. 21900.
- [26] Y. Liu and Z. Sun, "S-index: A new measure to evaluate the scientific impact of scholars," in *Proc. 3rd Int. Conf. Adv. Electron. Mater., Comput. Softw. Eng. (AEMCSE)*, Apr. 2020, pp. 529–535.
- [27] G. Mustafa, M. Usman, M. T. Afzal, A. Shahid, and A. Koubaa, "A comprehensive evaluation of metadata-based features to classify research paper's topics," *IEEE Access*, vol. 9, pp. 133500–133509, 2021.
- [28] M. Usman, G. Mustafa, and M. T. Afzal, "Ranking of author assessment parameters using logistic regression," *Scientometrics*, vol. 126, no. 1, pp. 335–353, Jan. 2021.
- [29] R. Liang and X. Jiang, "Scientific ranking over heterogeneous academic hypernetwork," in *Proc. AAAI Conf. Artif. Intell.*, 2016, vol. 30, no. 1, pp. 20–26.
- [30] M. Amer and M. T. Afzal, "Evaluation of h-index and its qualitative and quantitative variants in neuroscience," *Scientometrics*, vol. 121, no. 2, pp. 653–673, Nov. 2019.
- [31] S. M. A. H. Shah, A. Ullah, J. Iqbal, S. Bourouis, S. S. Ullah, S. Hussain, M. Q. Khan, Y. A. Shah, and G. Mustafa, "Classifying and localizing abnormalities in brain MRI using channel attention based semi-Bayesian ensemble voting mechanism and convolutional auto-encoder," *IEEE Access*, vol. 11, pp. 75528–75545, 2023.
- [32] L. Smolinsky and A. Lercher, "Citation rates in mathematics: A study of variation by subdiscipline," *Scientometrics*, vol. 91, no. 3, pp. 911–924, Jun. 2012.
- [33] M. Schreiber, "Self-citation corrections for the Hirsch index," *Europhys. Lett.*, vol. 78, no. 3, p. 30002, May 2007.

- [34] Q. L. Burrell, "Hirsch's  $h$ -index: A stochastic model," *J. Informetrics*, vol. 1, no. 1, pp. 16–25, Jan. 2007.
- [35] K. R. Dienes, "Completing  $h$ ," *J. Informetrics*, vol. 9, no. 2, pp. 385–397, 2015.
- [36] M. Schreiber, "A modification of the  $h$ -index: The  $h_m$ -index accounts for multi-authored manuscripts," *J. Informetrics*, vol. 2, no. 3, pp. 211–216, Jul. 2008.
- [37] 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, Apr. 2019.
- [38] R. Ghani, F. Qayyum, M. T. Afzal, and H. Maurer, "Comprehensive evaluation of  $h$ -index and its extensions in the domain of mathematics," *Scientometrics*, vol. 118, no. 3, pp. 809–822, Mar. 2019.
- [39] C. Moreira, P. Calado, and B. Martins, "Learning to rank academic experts in the DBLP dataset," *Expert Syst.*, vol. 32, no. 4, pp. 477–493, Aug. 2015.
- [40] R. Dillon, "The  $u$ -index: An alternative metric to measure academic impact," in *Proc. 12th Int. Conf. Virtual Campus (JICV)*, Sep. 2022, pp. 1–3.
- [41] E. Dunne and K. Hulek, "Mathematics subject classification 2020," *Nat. Amer. Math. Soc.*, vol. 67, no. 3, pp. 410–411, 2020.
- [42] G. Halevi, H. Moed, and J. Bar-Ilan, "Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the literature," *J. Informetrics*, vol. 11, no. 3, pp. 823–834, Aug. 2017.
- [43] G. Mustafa, A. Rauf, A. S. Al-Shamayleh, M. Sulaiman, W. Alrawagfeh, M. T. Afzal, and A. Akhunzada, "Optimizing document classification: Unleashing the power of genetic algorithms," *IEEE Access*, vol. 11, pp. 83136–83149, 2023.
- [44] S. Bhattacharya, A. Banerjee, A. Mazumder, and S. Nandi, "Impact of author indexing from the co-authorship relation," in *Proc. Int. Interdiscipl. Conf. Math., Eng. Sci. (MESICON)*, Nov. 2022, pp. 1–6.
- [45] N. S. Maurya, S. Kushwah, S. Kushwaha, A. Chawade, and A. Mani, "Prognostic model development for classification of colorectal adenocarcinoma by using machine learning model based on feature selection technique Boruta," *Sci. Rep.*, vol. 13, no. 1, p. 6413, Apr. 2023.
- [46] M. B. Kursa and W. R. Rudnicki, "The all relevant feature selection using random forest," 2011, *arXiv:1106.5112*.
- [47] M. B. Kursa and W. R. Rudnicki, "Feature selection with the Boruta package," *J. Stat. Softw.*, vol. 36, no. 11, pp. 1–13, 2010.
- [48] J. Lötsch, A. Ultsch, B. Mayer, and D. Kringsel, "Artificial intelligence and machine learning in pain research: A data scientometric analysis," *PAIN Rep.*, vol. 7, no. 6, p. e1044, Nov. 2022.
- [49] K. Nguyen and W. Chen, "DEM- and GIS-based analysis of soil erosion depth using machine learning," *ISPRS Int. J. Geo-Inf.*, vol. 10, no. 7, p. 452, Jul. 2021.



**LI WANG** is currently a Professor and a Ph.D. Advisor with the College of Data Science, Taiyuan University of Technology. She has undertaken more than 30 projects and published more than 100 academic articles. Her research interests include big data computing, machine learning, and artificial intelligence. She is a Senior Member of ACM, CCF, CCF Artificial Intelligence and Pattern Recognition Committee, and CCF Collaborative Computing Committee; a Standing Committee Member of the CCF Big Data Expert Committee; and a member of the Service Intelligence Committee of China AI Society. She was the program committee of many conferences and hosted and served for many conferences..



**AHMAD SAMI AL-SHAMAYLEH** received the master's degree in information systems from The University of Jordan, Jordan, in 2014, and the Ph.D. degree in artificial intelligence from the University of Malaya, Malaysia, in 2020. He is currently an Assistant Professor with the Faculty of Information Technology, Al-Ahliyya Amman University, Jordan. His research interests include artificial intelligence, human-computer interaction, the IoT, Arabic NLP, Arabic sign language recognition, language resources production, the design and evaluation of interactive applications for handicapped people, multimodality, and software engineering.



**MUHAMMAD TANVIR AFZAL** (Member, IEEE) received the M.Sc. degree in computer science from Quaid-i-Azam University, Islamabad, Pakistan, and the Ph.D. degree (Hons.) in computer science from the Graz University of Technology, Austria. He was associated with academia and industry at various levels for the last 20 years. He was with the Namal Institute Mianwali, as the Director and a Professor. He was with the Capital University of Science and Technology, Islamabad, as a Professor, an Associate Professor, and an Assistant Professor of CS. Furthermore, he was with NESCOM, COMSATS University Islamabad, and JinTech Islamabad. He is currently the Director and a Professor with the Shifa School of Computing and Director Campus at one of the largest campus of Shifa Tameer-e-Millat University, Islamabad. He has authored more than 125 research articles, including over 50 published articles in impact factor leading journals in the field of data science, information retrieval and visualization, semantics, digital libraries, artificial intelligence, and scientometrics. He has authored two books and has edited two books in computer science. His cumulative impact factor is more than 123, with citations over 1,300. He played pivotal role in making collaborations between MAJU-JUCS, MAJU-IICM, and TUG-UNIMAS. He has conducted more than 100 curricular, co-curricular, and extra-curricular activities in the last five years, including seminars, workshops, national competitions (ExcITeCup); and invited international and national speakers from Google, Oracle, IICM, IFIS, and SEGA Europe. Under his supervision, more than 70 postgraduate students (M.S. and Ph.D.) have defended their research theses successfully and a number of Ph.D. and M.S. students are pursuing their research with him. He served as a primary trainer and the program director at a national level training for a public sector organization in Pakistan on human factors engineering and conducted the training of over 150 hours for experts from the industry. He remained primary trainer for data science trainings in leading industrial organizations. He received the Gold Medal during the M.Sc. study. He was a recipient of multiple international research fundings. He served as the Ph.D. symposium chair, the session chair, the finance chair, a committee member, and an editor for several IEEE, ACM, Springer, Elsevier international conferences and journals. He is serving as the Editor-in-Chief for reputed impact factor journal, such as *Journal of Universal Computer Science*.



**BILAL AHMED** (Member, IEEE) received the master's degree in Pakistan. He is currently pursuing the Ph.D. degree with the Taiyuan University of Technology, China. He was associated with academia and industry for the last four years. In his academic career, he taught different computer science courses, such as data science, tools and techniques for data analytics, also courses on various programming languages, such as Python and C++. His research interests include machine learning, deep learning, data science, scientometrics, and recommendation systems.



**GHULAM MUSTAFA** received the B.S. degree in software engineering from COMSATS, Abbottabad, in 2017, and the M.S. degree (Hons.) in computer science from the Capital University of Science and Technology, Islamabad. He is currently pursuing the Ph.D. degree in computer science with the University of Engineering and Technology, Taxila, Pakistan. He was associated with academia and industry for the last six years. He is also a Senior Lecturer with the Faculty of Computing, Shifa Tameer-e-Millat University, Islamabad. Previously, he was with the CS Department, Capital University of Science and Technology, as an Associate and a Junior Lecturer, for five years. Furthermore, he was with A&F Solution Software House, as a Web Frontend Designer and a Backend Developer. Moreover, with academia, he is also an active freelancer for the last four years, doing projects in different languages such as, python, java, and C++. In his academic career, he has taught different computer science labs, such as Introduction to Programming Lab (C++), Object Oriented Programming Lab (C++), Advanced Computer Programming Lab (JAVA), Database System Lab, Data Structure Lab (C++), Computer Communication and Network Lab (CNN), and Web Development Lab, while serving as a Junior Lecturer. He has taught different computer science courses, such as introduction to programming, object oriented programming, discrete structure, theory of automata, and software engineering, while serving as a Senior Lecturer and an Associate Lecturer. During his B.S. degree, he was awarded two Gold Medals on his first position in Abbottabad Campus, and got first position in all seven campuses of CUI, in 2017, and the Gold medal due to his excellent academic performance in his entire degree duration, during the M.S. study.



**WAGDI ALRAWAGFEH** is received the Ph.D. degree in computer science (multi-agent systems) from the Memorial University of Newfoundland, Canada. He has over ten years of experience working with environmental aspects directly related to education, programming, and information technology.



**ADNAN AKHUNZADA** (Senior Member, IEEE) is an accomplished cybersecurity specialist and consultant, boasting extensive industrial experience and expertise. He has made significant contributions to the cybersecurity landscape through impactful published research, successful commercial products, and holding U.S. patents. He has advised some of the largest ICT companies worldwide, effectively securing multi-million-dollar projects. As a recognized authority in the field, his expertise spans the design of innovative SIEM systems, IDS, IPS, threat intelligence platforms, secure protocols, AI for cybersecurity, secure future internet, adversarial, and privacy preserving machine learning. With a combination of proven industry success, technical prowess, and a dedication to advancing cybersecurity, he remains at the forefront of the field, continuously shaping its future. With over a decade of experience in the field, he is highly regarded as a Professional Member of ACM.

• • •