Knowledge and Information Systems (2024) 66:5203–5238
https://doi.org/10.1007/s10115-024-02119-1

REGULAR PAPER

GK index: bridging Gf and K indices for comprehensive author evaluation

Ghulam Mustafa[1] · Abid Rauf[1] · Muhammad Tanvir Afzal[2]

Received: 25 December 2023 / Revised: 28 March 2024 / Accepted: 1 April 2024 /
Published online: 25 April 2024

Abstract

The task of accurately predicting scientific impact and ranking the researcher based on impact has emerged as a crucial research challenge, captivating the interest of scholars across diverse domains. This task holds immense importance in enhancing research efficiency, aiding decision-making processes, and facilitating scientific evaluations. For this, the scientific community has put forth a wide array of parameters to identify the most influential researchers. These include citation count, total publication count, hybrid methodologies, the h-index, and also its extended or modified versions. But still, there is a lack of consensus on a single optimal parameter for identifying the most influential author. In this study, we introduce a novel index derived from learning hidden patterns within the mathematics field dataset, comprising data from 1050 researchers evenly split between awardees and non-awardees. Initially, we ranked selected parameters by assessing their values for individual researchers, identifying the top five parameters that most frequently placed awardees within the top 100 records. Additionally, we employed deep learning techniques to identify the top five influential parameters from the initially selected set. Subsequently, we evaluated the disjointness between the results produced by these parameters. To further refine our analysis, we assessed seven different statistical models for combining the top disjoint pair to retain the maximum properties of both parameters. The study's findings revealed that the gf and k indices exhibited a 0.96 percent disjointness ratio, establishing them as the highest disjoint pair. Moreover, the geometric mean demonstrated a 0.87 percent average impact in retaining the properties of the top disjoint pair, surpassing the other seven models. As a result of this study, we propose a new index obtained by taking the geometric mean of the top disjoint pair

Ghulam Mustafa, Abid Rauf and Muhammad Tanvir Afzal have equally contributed to this work.

✉ Ghulam Mustafa
ghulam.mustafa2@students.uettaxila.edu.pk

Abid Rauf
abid.rauf@uettaxila.edu.pk

Muhammad Tanvir Afzal
director.ssc@stmu.edu.pk

1 Department of Computer Science, University of Engineering and Technology, Taxila 47080, Pakistan

2 Department of Computing, Shifa Tameer-e-Millat University, Islamabad 44000, Pakistan

which increase the result by 12% as compared to existing best performing individual index performance.

Keywords Author assessment parameters · Parameter ranking · Multi-layer perceptron (MLP) · Variants of h-index · Mathematics domain datasets

1 Introduction

Ranking researchers and identifying influential authors within the scientific community is a crucial research endeavor with far-reaching implications. These implications include recognizing individuals deserving of tenure, securing funding, obtaining project opportunities, or awarding recognition to impactful researchers [1–4]. Moreover, rankings play a pivotal role in evaluating academic output and accomplishments. Raheel et al. [5] emphasize that a fair and comprehensive ranking system that considers educational qualifications could simplify the selection of keynote speakers for conference organizers. Additionally, it can empower students to make informed decisions while selecting a professor to supervise their research, based on their academic standing and capabilities. Therefore, there is entire need to establish a researcher's index that can effectively rank individuals within a scientific community.
Henceforth, numerous researchers have worked to propose author assessment parameters over the last two decades. More than 70 different author assessment parameters have been suggested by various researchers and extensively debated in the literature [6, 7]. These parameters employ distinct criteria to evaluate impact, encompassing both quantitative and qualitative aspects, with some displaying hybrid characteristics. Initially, researchers were assessed based on metrics like publication quantity [8] and citation counts [9]. However, these metrics alone do not capture the entirety of a researcher's impact. It is not necessarily the case that a large number of publications by a researcher reflects the quality of their work, as some authors may begin publishing articles in journals with low impact factors merely to increase their total publication count [10]. Similarly, high citation counts may not reliably indicate quality or lasting influence due to practices like self-citation and publishing survey papers to increase citations [11]. The h-index was introduced to address these concerns, offering a measure of both productivity and impact [12]. Nonetheless, it too has limitations, such as overlooking total citations of highly cited publications and the collaborative aspect of research [13]. Despite the plethora of ranking parameters, a strong consensus on a primary measure for ranking researchers has not been reached, prompting ongoing debate and further research. Furthermore, besides proposing new indices, several studies across different domains have evaluated parameters, uncovering varying optimal parameters for different datasets and no consistent ranking across different studies exists [4, 11, 14–16]. However, these studies have often focused on a limited number of parameters, constraining a comprehensive understanding of their overall impact.
After conducting a comprehensive analysis of existing literature, several issues have emerged. Firstly, new indices are often tested in hypothetical or fictional scenarios rather than being validated on comprehensive datasets within specific domains [17–22]. Secondly, existing approaches fail to evaluate indices against a common yardstick or gold standard, hindering effective evaluation. Thirdly, while different indices are based on common parameters such as publication, citation, and co-authorship, they behave differently when their formulas are computed or proposed. Consequently, there is a lack of studies evaluating each index's ability to return awardees in a disjointed manner. The disjointness value reaches hundred between

two parameters when the first parameter returns the maximum number of unique awardees, or when it intersects with the result set of the second parameter, resulting in an empty set of awardees.

To address the issues mentioned above, this study tends to adopt a comprehensive approach. Hence, the main contributions of this study are as follows:

- Conducted a thorough analysis of author assessment parameters using a comprehensive dataset from the mathematics domain, previously utilized in various studies [4, 23].
- Meticulously selected a large number of existing parameters, totaling sixty-four, a scope unprecedented in existing literature.
- Calculated the disjointness ratio between all possible pairs of top-performing indices.
- Employed a comprehensive approach to evaluate indices on a gold standard dataset, identifying the top pair and combining them using different statistical models.
- Proposed a new index based on the top-performing statistical model.

After analyzing the mentioned contributions, we aim on implementing the methodology referred in the paragraph ahead. Initially, we chose datasets of mathematics domain containing instances of both awardees and non-awardees. Furthermore, for the collection of existing indices from literature, we established specific criteria and created four categories: primitive parameters, parameters based on publication and citation counts, parameters based on publication age, and parameters based on author counts. From the literature, we selected 64 parameters and categorized them accordingly. To evaluate these parameters and identify the most influential ones, our methodology employed two approaches: ranking authors based on calculated index values and applying deep learning techniques to rank parameters. The primary objective of both approaches was to identify the parameters that resulted in the highest number of awardees within their top 100 records. Subsequently, to mitigate biases, we selected the top 5 parameters from both approaches (10 in total) and generated all possible combinations (45 in total), analyzing the disjointness between their results. Additionally, we selected the top disjoint pair (gf index and k index) and combined them using seven different statistical models to determine the most effective model (geometric mean) that retained the maximum characteristics of this pair. Furthermore, in proposing a new index, we integrated weight factors with indices in the formula to signify their importance relative to others. These weights were calculated based on the position of the index in both ranking approaches. The newly combined proposed index increased the number of awardees in the top 100 records by 12% compared to the number of best performing individual index awardees in the top 100 records.

The rest of this paper is structured as follows: In Sect. 2, we provide a concise overview of relevant studies. Section 3 outlines the approach we used to calculate, evaluate the indices, and propose a new index. Section 4 discusses the outcomes of our study, and Sect. 5 presents our concluding remarks.

2 Literature

Evaluating the scientific productivity of authors holds significant importance across various research-related domains. It plays a pivotal role in making well-informed decisions, selecting candidates for scientific awards, matching researchers with suitable research projects, facilitating promotions, granting tenures to deserving individuals, and awarding contracts to experts in the field [24]. In the existing literature, more than 70 parameters have been proposed over the past two decades to assess and rank researchers based on their research contributions [25]. Initially, the parameters used to evaluate a scientist's research is the total number of publication [10]. However, this metric falls short in capturing a comprehensive analysis of the impact and quality of an researcher's work, as some authors start to publishing articles in journals with low impact factors merely to boost their total publication count. In response to this limitation, citation count was proposed [9]; nevertheless, it carries its own set of drawbacks. For instance, it can take time for newly published papers to accumulate citations, which can be a disadvantage for emerging authors, even if they possess a substantial publication count. Moreover, citation count may not always provide an accurate gauge of publication quality, as researchers may cite papers by others primarily for critique rather than endorsement. Furthermore, researchers sometimes engage in self-citation, citing their own work to inflate their citation numbers.

In response to the limitations inherent in relying solely on publication & citation counts, Hirsh, 2005 [12] introduced the h-index, a metric that swiftly gained popularity and became the predominant parameter for assessing an author's research output. However, the h-index itself is not without its shortcomings. It is primarily rooted in long-term observations, which means that older authors tend to have larger h-indices than their newer counterparts, and it is influenced by the specific field of research. Additionally, the increase in citations for an author's core articles does not significantly impact the overall h-index. To overcome these limitations, numerous new indices and variations of the h-index have been put forward, including the g index [17], t-index [18], AR index [19], e index [20], P index [21], A index [22], m-quotient [26], contemporary h-index [26], f-index [27], Wu-index [28], and q2-index [10] and so on. One of the primary limitations of these techniques are, when introduced in the literature, it has often tested in hypothetical or fictional case scenarios [29].

Numerous studies have been carried out to assess various research indices across diverse datasets. For instance, Kolmulski et. al. [30] was undertaken a study to explore the correlation between the h-index and the h(2) index, utilizing data from 19 professors associated with the chemistry department at the University of Poland. The study revealed a strong correlation between these two indices, indicating that their outcomes closely align and produce similar results. Likewise, Vaan Raan [31] conducted a study comparing the h-index and its variants, employing data collected from a thorough assessment of 147 research groups within the chemistry department in the Netherlands. Unlike many studies that primarily focus on individual performance, this study centered on assessing the performance of research groups. Notably, the study took into account only three-year citations instead of considering an author's entire lifetime citations, aiming to calculate the correlation among the h-index and its various adaptations within this specific context. In their study, Jin et al. [32] introduced a novel approach to evaluate authors by integrating multiple indices. This method employs a combination of two indices one serving as a quantitative measure and the other as a qualitative measure. The pairs of indices considered in the study included metrics like the h-index, r-index and ar-index. The results of this investigation demonstrated that the combination of the h-index and ar-index proved to be effective indicators for assessing research performance. Mane et. al. [33] carried out a study to assess the g-index in comparison with other indices, including the h-index, a-index, and r-index. This analysis was conducted using data from a group of 26 physicists affiliated with the Physics Department at Chemnitz University of Technology. The results of the study led Schreiber to conclude that the g-index serves as a more suitable metric for gauging the overall impact of a scientist's publications when compared to the h-index. In Xiao et. al. [34] conducted an extensive evaluation of 29 different h-index variants. Their study involved calculating the correlations between these variants and two benchmark indices, the h-index and Wu index. Notably, the research aimed to identify the level of correlation each variant had with these benchmark indices. What they found was that indices highly correlated with the h-index tended to exhibit lower correlations with the Wu index, and vice versa. Furthermore, the study reported that the highly correlated indices demonstrated only a marginal improvement over the h-index or Wu index. One of another interesting study in which the Xindong [35] proposed w index. In this study the author addressing the limitations of existing single-number metrics like the h-index and q-most-citations, the w-index considers the significance of representative publications by weighting the citations of the researcher's top publications based on authorship position. This approach aims to provide a more nuanced assessment of research impact beyond mere citation counts. Furthermore, the paper acknowledges the challenges associated with self-citations but emphasizes their minimal impact based on empirical analysis. In cases where calculating the w-index may be impractical, the paper suggests prioritizing the three most cited first-authored publications alongside the h-index. Overall, the w-index aligns with established principles of recognizing the importance of highly cited publications while also acknowledging that citations alone should not be the sole criterion for evaluating research impact.

Furthermore, Ayaz et. al. [36] conducted a study in which they evaluated the complete-h index using a benchmark dataset consisting of awardees from mathematical scientific societies. Their findings revealed that the complete-h index outperformed other indices within this context. Similarly, Raheel et. al. [5] conducted an evaluation of the h-index and its variants based on publication age and citation intensity within the field of civil engineering. The results of this study indicated that the Wu index performed better than other indices. Ain et al. [36] evaluated the h-index and its variants using a mathematics domain dataset. They reported that the fraction count on paper parameter outperformed others, identifying 55 percent of awardees within the top 10 records. Additionally, Ameer et. al. 2019 [37], in the same year, assessed quantitative parameters using a dataset from neuroscience societies. Their study highlighted the effectiveness of the R-index and hg-index in ranking awardees at the top positions. In another study, Ain et. al. [16] aimed to evaluate quantitative parameters for researchers using a mathematics domain dataset. They sought to determine the correlation between the h-index and its variants for researchers and ranked these parameters based on their association with award-winning researchers. One of the researchers, Salman et. al. [38] evaluated indices based on multi-authorship, employing a civil engineering dataset. They reported that the gf index outperformed all other indices. However, these studies had a limitation in that they aimed to establish a link between the h-index or its variants and researchers who had received awards before these parameters were proposed. Consequently, it is possible that these awardees were not reliant on the h-index or its variants, and any correlations found could have been coincidental. To address this limitation, Usman et al. [39] conducted a study in which they evaluated the h-index and its variants using data from researchers in the field of civil engineering. They did not select researchers randomly; instead, they chose both awardees and non-awardees from the same time period, with a focus on those who received awards after 2005. This approach allowed for more conclusive results regarding the parameters crucial for identifying award winners. Moreover, in a recent study, Alshdadi et. al. [40] proposed a rule for the scientific community utilizing deep learning models. They utilized datasets from various domains, including civil engineering, mathematics, and neuroscience, incorporating data from 500 researchers in each field. Their findings demonstrated that these rules achieved an accuracy rate of up to 70%. Finally, Mustafa et. al. [11, 14] conducted two recent studies. In the first, they evaluated publication and citation count-based parameters, discovering that the normalized h-index outperformed all other parameters. In the second study, they assessed publication age-based parameters, revealing that the Ar index outperformed the remaining indices. For experimental purposes, they utilized a mathematics domain dataset. One of

another author, Ahmed et. al. [4] used the similar dataset and evaluated the author count based parameter. They reported that hf index outperformed all other indices.
After conducting an extensive review of the existing literature, we embarked on a detailed examination of numerous proposed techniques, complemented by comprehensive evaluation surveys of these parameters. We observed that, initially, the assessment was heavily reliant on publications and citations counts, but this approach underwent a significant transformation in the subsequent decade. The focus shifted toward utilizing various forms of the h-index, often without accounting for the specific limitations or contextual considerations of these metrics. The introduction of novel methodologies frequently occurred in unconventional ways, and they were tested across diverse datasets. This divergence in data sources and validation methods complicates the assessment of the significance of individual indices. Hence, there is a requirement for a study that evaluates a vast array of parameters using a single-domain dataset and identifies the top-performing parameter through straightforward ranking. Additionally, we should employ machine learning and deep learning models to rank these parameters systematically and pinpoint the most influential ones that consistently place awardees in the top 100 records. Furthermore, there is also a necessity for the development of a novel index that can yield a greater number of awardees compared to the currently top-performing indices.

3 Methodology

After conducting an in-depth literature review, we turn our attention to proposed a new index for researcher's assessment for mathematics fields. Figure 1 presents the architecture diagram, encompassing various stages, including: (i) dataset collection, (ii) parameter calculation, (iii) simple ranking, (iv) ranking with deep learning, (v) disjointness analysis (vi) statistical analysis and, (vii) proposed index. In the subsequent sections, we will delve into the detailed explanation of these Steps.

3.1 Dataset collection

To conduct our experiments, we needed a dataset belonging to a particular domain. For this study, we chose the field of Mathematics as our primary focus, due to its rich historical significance and substantial contributions to research. This selection makes it the ideal domain for assessing the proposed methodology. Furthermore, within the field of Mathematics, various esteemed scientific societies annually recognize outstanding researchers based on the impact of their work. Given the breadth and depth of this domain, numerous previous studies have also used it [4, 11, 16]. Within this domain, we opted for the dataset utilized by [11] and [4] in their respective research studies. This dataset contains 1050 records, providing comprehensive information about both awardees and non-awardees. Specifically, the dataset comprises 525 non-awardees and 525 awardees. Mustafa et. al. [11] assembled a list of 30 prestigious international awards of significant importance within the mathematical community for the awardees' data. These awards represent esteemed achievements for mathematicians and researchers and are conferred by renowned mathematical societies such as the London Mathematical Society (LMS), International Mathematical Union (IMU), Norwegian Academy of Science and Letters (NASL), and the American Mathematical Society (AMS). The proportion of data pertaining to non-awardees in the dataset was derived from the datasets used by Ain et al. [16] and Ghani et al. [41]. However, it is worth noting that the original

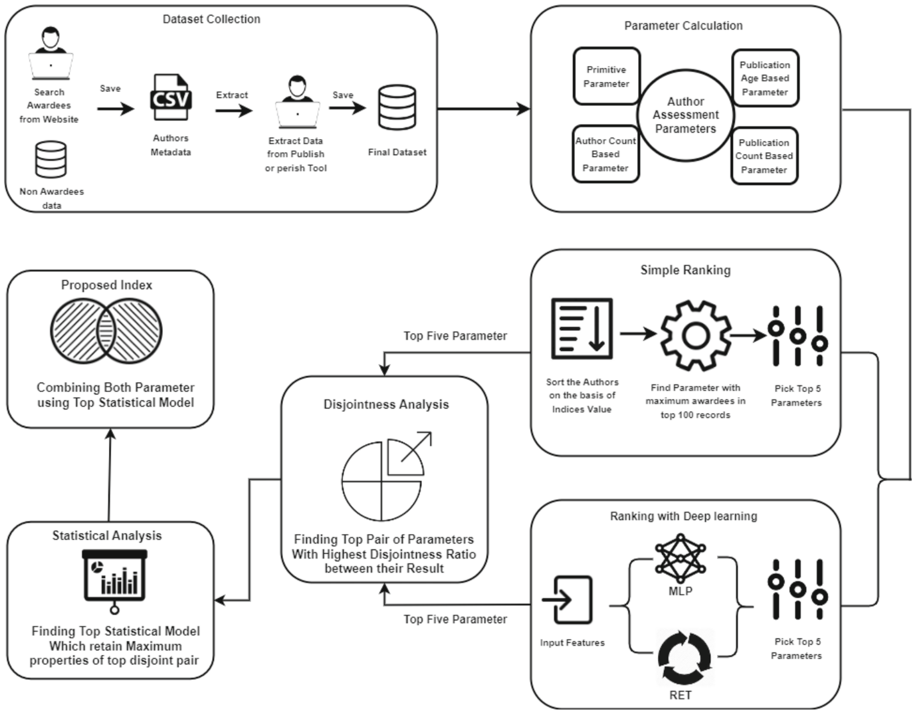

Fig. 1 Architecture diagram of proposed methodology

datasets employed by Ain et al. [16] and Ghani et al. [41] contained only a limited number of entries for awardees, extending only up until 2013. To address this limitation, Mustafa et al. [11] expanded the dataset by acquiring updated information on awardees, spanning up to the year 2023. A comprehensive overview of the dataset's statistics is provided in Table 1. To gather data on awardees, they have visited various society websites within the field of mathematics of last three decades, collecting names and corresponding award years for researchers. For extraction of researcher data, they have used the Publish or Perish platform, implementing a 'hold on' strategy that allowed us to collect researchers' data even before their award-receiving years. This tool employed a sophisticated algorithm to retrieve metadata from Google Scholar. To ensure fairness and balance within the dataset, they have gathered data on non-awardees in a proportion equivalent to the number of awardees for each year. For example, if there were 15 awardees in 1991, we employed similar techniques to collect data from the 15 non-awardees prior to 1991. Before embarking on any analysis or evaluation, it was imperative to conduct a meticulous data cleansing process, particularly for data obtained from sources like Google Scholar. This process was aimed at eliminating inaccuracies or irrelevant information, often referred to as 'noise,' which could potentially compromise the accuracy of our results. The data cleansing procedure consisted of several crucial steps, including data accuracy verification and the removal of duplicate entries. Within this extensive research dataset, two pivotal processes were executed to enhance the data's quality and relevance. Initially, a filter was applied to ensure the consistency of each research article within the mathematical field, eliminating any extraneous content. This step was instrumental in refining the dataset, focusing solely on the relevant domain. Subsequently, numerous

Table 1 Dataset statistics

Researchers metadata	Count
Total authors records	1050
Total awardees	525
Total non-awardees	525
Total citation	14,370,007
Total publication	204,896

researchers have highlighted the necessity to disambiguate author data collected from Google Scholar [42]. In this study, we analyze two samples of data: awardees and non-awardees. The full names of awardees are extracted from society websites, totaling 525 individuals. Interestingly, we found no authors requiring disambiguation among the awardees. Conversely, for non-awardees, we gathered data from previously published datasets utilized by Ain et al. [16] and Ghani et al. [41]. During the author disambiguation process, two primary cases emerged. Case 1 involves identical last and first names, requiring verification of their distinctiveness. Case 2 pertains to identical last names but different first names, which often necessitate evaluation. We employed methodologies documented in existing literature [16, 36] to address these cases. Among the 525 non-awardees, no authors shared both first and last names, rendering Case 1 inapplicable. However, 43 samples within this group shared identical last names, becoming candidates for disambiguation. Remarkably, our analysis revealed that 26 out of 43 authors shared the same last names but different first names, representing distinct individuals. Conversely, 17 out of 43 authors shared identical last names and different first names, yet they were variations of the same individual. To ensure balance, we augmented the dataset with additional unique authors. Furthermore, the "publish or perish" tool provided a valuable mechanism for self-checking disambiguation during data collection. Its smart search feature allows for the inclusion of multiple variations of an author's name, facilitating more accurate data collection. The link provides access to the complete dataset comprising 1050 instances. [1]

3.2 Calculation of indices

In this section, we perform calculations for the sixty-four author assessment parameters, categorized into four distinct groups, using the dataset we've gathered. These groups are primitive parameters, publication and citation count-based parameters, Author count-based parameters and publication age-based parameters. The categories and their corresponding indices and its formulas along with detailed explanations are presented in Table 4 (located in Appendix A). Moreover, we have also make a small application which is used for the extraction of researcher data by using the google scholar profile link. Furthermore, by using the extracted data against author the application calculates all of the 64 parameters against that author. The main interface of the application is shown in Fig. 2; the code of the application is provided on the mentioned link. [2]

[1] https://github.com/ghulammustafacomsat/Mathematics_dataset.

[2] https://github.com/ghulammustafacomsat/dataextraction.

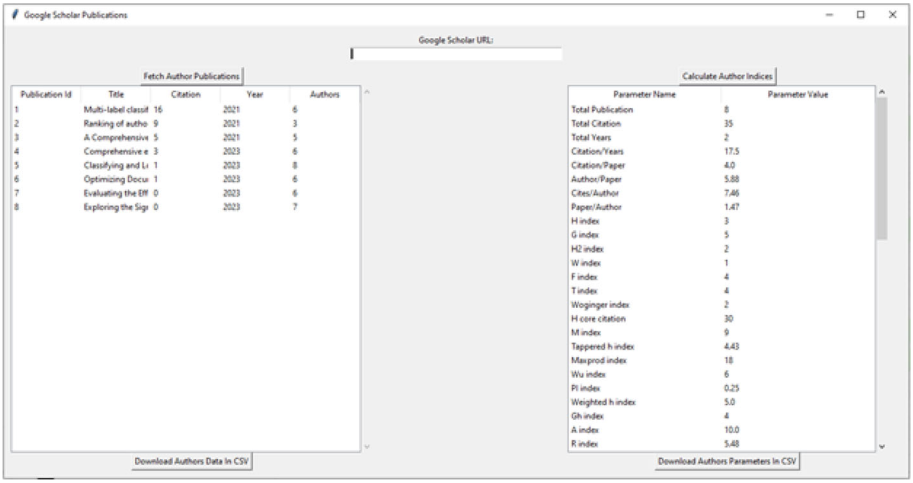

Fig. 2 Application interface

3.3 Simple ranking

In a simple ranking approach, we begin by creating distinct lists of authors for each index. Subsequently, we sort the authors within each index list based on their index values. Following this, we determine the number of awardees present in the top 100 records of each ranked list for every parameter. This provides insight into the significance of each index in determining author rankings. Finally, we select the top 5 parameters that yield the highest number of awardees in the top 100 records. Algorithm 1 represents the way to extract the top 5 parameters which return the highest number of awardees.

3.4 Ranking with deep learning

In the realm of machine learning, the task of feature ranking holds significant importance, playing a pivotal role in numerous tasks such as dimensionality reduction, interpretability enhancement, overfitting mitigation, prediction optimization, and feature selection. In this study, we employ a multilayer perceptron classifier (MLP) coupled with a recursive elimination technique to identify the importance score of features. The MLP is a neural network with multiple hidden layers, utilizing rectified linear unit (ReLU) activation for the hidden layers and softmax activation for the output layer. A critical aspect of the study is the use of a loss function to gauge the disparity between predicted and actual values, with iterative backpropagation to optimize the model by adjusting weights and biases. To counteract overfitting, batch normalization is thoughtfully integrated after each hidden layer. The model comprises 10 hidden layers, each with 10 neurons and is trained using the Adam optimization algorithm with a learning rate of 0.0003. Additionally, an early stopping technique is implemented to halt training when signs of overfitting emerge, preserving the model's generalization capabilities. In addition, our classification model employs a recursive elimination technique (RET) with some modifications for calculating the importance scores. This technique is widely used to identify relevant features that significantly contribute to the model's performance [43, 44]. Moreover, the reason for choosing MLP with RET lies in its simplicity and ability to strike

Algorithm 1 Simple Ranking of Authors (Return Top 5 Parameters)

```
Indices_value ← Authors_Parameters_data
Researcher_label ← Authors_label(Awardees, Non − Awardees)
Index_Values_list ← []
Label_list ← []
Index_name ← []
awardees_count ← []
i ← 1
while i < 64 do
    for j in 1 to len(totalauthors) do
        Index_Values_list.append(Indices_value[j])
        Label_list.append(Researcher_label[j])
    end for
    Index_Values_list, Label_list ← Sort(Index_Values_list, Label_list)
    count ← 0
    for K ← 1 to 100 do
        if Researcher_label[K] == 1 then
            count ← count + 1
        else
            Continue
        end if
    end for
    Index_name.append(Indices_value[0])
    awardees_count.append(count)
    i ← i + 1
end while
Index_name, awardees_count ← Sort(Index_name, awardees_count)
return Index_name.head(5)
```

a balance between model complexity and performance, especially when dealing with small datasets. When datasets are not extensive or lack the complexity necessitating sophisticated deep learning architectures, deep learning techniques frequently run the risk of overfitting on smaller datasets if not properly regularized [45]. Furthermore, RET effectively reduces the dataset's dimensionality, enhancing model interpretability, efficiency, and generalization ability [29]. RET iteratively eliminates irrelevant or redundant features, focusing on a subset of features with the most significant impact on the model's performance. In this study, initially, the dataset was divided into two sets in an 80:20 ratio: the first for training and the second for testing. Within the training set, we further divided it into an 80:20 ratio. Here, 80% was allocated for actual model training, while 20% was reserved for validation during the training phase. Furthermore, the distribution of instances between these sets is balanced, indicating an equality in the number of different classes. The training dataset is used to train a multilayer perceptron classifier for classification purposes, as discussed earlier in this section. Subsequently, a validation and test sample is provided to the trained model, and the model predicts the class label for each sample. The accuracy achieved during this prediction stage serves as the baseline accuracy when all features are included. The next phase of the technique focuses on feature removal. One parameter is removed from the feature list, and the dataset is once again divided into training and test sets (similarly, as mentioned in the preceding statement). The multilayer perceptron classifier is then trained using the updated feature set. After training, a test sample is used to predict the class label, and the accuracy is recorded. The new accuracy obtained is subtracted from the baseline accuracy, resulting in a subtraction result that serves as the importance score for the removed feature. This process is repeated for each parameter with at least five different epoch phases, yielding an importance score. Equation (1) represents the calculation of the importance score.

$$F_{IS} = \frac{1}{5}\sum_{i+=20}^{100}(BLA_i - WOPA_i) \tag{1}$$

where F_{IS} represents feature important score, i represents the number of epochs, BLA_i represents the baseline accuracy against the i^{th} phase, and $WOPA_i$ represents the without-parameter accuracy of the i^{th} phase.
The entire process was repeated for each parameter in the dataset. After iterating through all the features, the algorithm generates two lists. The first list contains the names of the features, while the second list contains their corresponding importance scores. Based on these importance scores, the parameters are sorted, resulting in parameter ranking. From the resulted ranking, we have picked the top 5 most influential parameters. Algorithm 2 represents the way to extract the top 5 parameters which returns a highest no of awardees while using deep learning model.

Algorithm 2 Ranking With Deep Learning Model

```
Input: Features with Class label                                    ▷ Features and its Class Label
Output: Return Top 5 Parameter with Importance Score
while i ≤ len(Features) do
    if i == 0 then
        X ← All Parameters Data                                     ▷ Load parameters data except class label
        y ← Class Label                                             ▷ Load class label of all records
        Xtrain, Xtest, ytrain, ytest ← Split(X, y, 0.20)            ▷ Split the data
        mlp ← BuildMLPClassifier()                                  ▷ Build Multilayer Perceptron
        mlp.fit(Xtrain, ytrain)                                     ▷ Model fit on data
        ypred ← mlp.predict(Xtest)                                  ▷ Class Label prediction
        baseAccuracy ← accuracyScore(ytest, ypred)                  ▷ Accuracy with all parameters
    else
        X ← X                                                       ▷ Assign data after removal of parameter
        y ← Class Label                                             ▷ Load class label of all records
        Xtrain, Xtest, ytrain, ytest ← Split(X, y, 0.20)            ▷ Split the data
        mlp ← BuildMLPClassifier()                                  ▷ Build Multilayer Perceptron
        mlp.fit(Xtrain, ytrain)                                     ▷ Model fit on data
        accuracy ← accuracyScore(ytest, ypred)                      ▷ Predict Accuracy after removal of parameter
        parameterName.append(X[i].name)  ▷ Append the Name of Parameters Which can Eliminate from Parameter List
        importanceScore.append(baseAccuracy-accuracy)  ▷ Importance Score of Parameters Which are Eliminating in iteration List
    end if
    X ← All Parameters Data             ▷ Load all data before removing index parameter in each iteration
    X.remove(i)                                                     ▷ Removing Index parameter
    i ← i + 1
end while
parameterName, ImportanceScore = Sort(parameterNameList, ImportanceScoreList)  ▷ Sort the Parameter Based on Important Score
return parameterName.head(5)                                        ▷ Return top 5 Parameter with Highest Score
```

3.5 Disjointness

In this study, the fundamental task at hand revolves around the examination and comprehension of disjointness among the top performing indices results. Disjointness, a critical aspect in many domains of studies, is central to discerning relationships, overlaps, and disparities within the top performing indices results. For this we have picked the top 5 parameters from both ranking processes. Subsequently, we have calculated the disjointness ratio between all the possible combination of the combined top 10 parameters. The disjointness ratio (DR) between parameters pair can be defined as:

$$DR = \frac{\Big(|P1| + |P2| - 2 * |P1 \cap P2|\Big)}{|P1| + |P2|} \tag{2}$$

In the above equation, the P1 represents first parameter of the pair and P2 represents the second parameter.

The disjointness ratio represents the extent to which both parameters yield results that do not overlap. For instance, if the first parameter returns 50 awardees within the top 100 records and the second parameter also returns 50 awardees within the same top 100 records, but there are 25 awardees that are common to both parameter results and 25 records that are unique to each parameter, then it means that there is a 50 percent disjointness between these two parameters. When the results returned by the parameters are entirely dissimilar, it signifies a disjointness ratio of 100 percent. Conversely, if the results returned by the parameters are entirely identical, the disjointness ratio is 0 percent.

Algorithm 3 Finding Disjointness Between Parameters

Input: Top 10 Parameters Data with Class label
Output: Return Parameters Combination with disjointness Ratio
while $i \leq$ len(Parameters) **do**
  $Author\_Name\_P1$, $Parameter\_1$, $Class\_label \leftarrow$ Sort($Author\_Name\_P1[i]$, $Parameter\_1[i]$, $Class\_label[i]$)
  while $k \leq 100$ **do**
    if $Class\_label[k] == 0$ **then**
      $parameter\_1\_awardees\_list.append(Author\_Name\_P1[i])$
    end if
  end while
  $j \leftarrow i + 1$
  while $j \leq$ len(Parameters) **do**
    $Author\_Name\_P2$, $Parameter\_2$, $Class\_label \leftarrow$ Sort($Author\_Name\_P2[j]$, $Parameter\_2[j]$, $Class\_label[j]$)
    while $n \leq 100$ **do**
      if $Class\_label[n] == 0$ **then**
        $parameter\_2\_awardees\_list.append(Author\_Name\_P2[i])$
      end if
    end while
    $Total\_Returned \leftarrow$ len($parameter\_1\_awardees\_list$) + len($parameter\_2\_awardees\_list$)
    $Common\_Returned \leftarrow$ Intersection($parameter\_1\_awardees\_list$, len ($parameter\_2\_awardees\_list$))
    $Disjointness \leftarrow (Total\_Returned - 2 \times Common\_Returned)/Total\_Returned$
    Right_in_CSVfile($parameter\_1$, $parameter\_2$, $Disjointness$)
    $j \leftarrow j + 1$
  end while
  $i \leftarrow i + 1$
end while

Table 2 Statistical methods

Method names	Formulas
Arithmetic mean	$Arithmetic = \frac{X_1+X_2+....+X_n}{n}$
Harmonic mean	$HarmonicMean = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}}$
Contra-harmonic mean	$Contra - HarmonicMean = \frac{(X_1^2)+(X_2^2)+\cdots+(X_n^2)}{(X_1+X_2+X_3...+X_n)}$
Geometric mean	$GeometricMean = (X_1 * X_2 * X_3 * ... * X_n)^{\frac{1}{n}}$
Logarithmic mean	$LogarithmicMean = (\frac{1}{n}) * (\sum_{i=1}^{n} log(X_i)))$
Root mean square	$RootMeanSquare = \sqrt{\frac{x_1^2+x_2^2+...+x_n^2}{n}}$ square=
Trigonometric mean	$TrigonometricMean = \frac{\prod_{i=1}^{n} sin(x_i)}{\prod_{i=1}^{n} x_i}$

3.6 Statistical analysis and proposed index

In light of our disjointness analysis results, we have identified and prioritized pairs characterized by the highest disjointness ratios. The presence of the highest disjointness ratio among these pairs serves as an indicator of dissimilarity in the returned awardee records. This phenomenon underscores that each parameter operates on a distinctive set of features, thereby resulting in divergent awardee selections. Consequently, this observation opens the door to utilizing the distinctive feature sets of both parameters. Such an approach could yield the creation of a novel index capable of returning awardees drawn from both parameter sets, thereby enriching our understanding of the underlying data.

For this we utilized the statistical models to combine value of these pairs. In general, statistical methods play a fundamental role in the analysis of data in various research domains (Tzenios, 2023). In this study, we employed top statistical techniques to gain deeper insights in these pairs and derive meaningful information. By harnessing the power of statistical analysis, we were able to systematically examine the data, identify patterns, quantify relationships, and make informed inferences. In this study, we aimed to combine top pairs using various statistical analysis methods. These methods include the arithmetic mean, contra-harmonic mean, geometric mean, harmonic mean, Lehmer mean, logarithmic mean, root mean square (RMS), and trigonometric mean. By employing these methods, we can obtain a comprehensive understanding of author rankings and assess their significance within the dataset. The calculations for these methods are presented in Table 2.

Furthermore, we employed the above sets of statistical methods (presented in Table 2) to analyze the top-ranked pair. By utilizing this list of statistical methods, we calculated the corresponding statistical method values for each pair. Subsequently, eight distinct lists were generated for each pair corresponding to each statistical method. Moreover, we compared these lists to discern the most influential statistical method for top ranked pair. Additionally, this analysis allowed us to identify the potential model for combining the features of top pair.

3.6.1 Validation of proposed index

In literature, authors employ various methods to validate their proposed indices. One common approach involves utilizing data from esteemed scientific academies such as the National

Academy of Sciences or the European Academy of Sciences. These datasets serve as benchmarks for evaluating the efficacy of the proposed indices. Another prevalent method involves the creation of hypothetical scenarios, wherein researchers simulate data from renowned professors across various educational institutions. This form of validation aims to assess the index's performance within specific organizational contexts, including universities and colleges. However, our analysis has identified an alternative validation approach from literature that appears more realistic. This method entails evaluating the proposed indices against prestigious awards received by individuals worldwide. These awards, spanning diverse fields and institutions, represent a global recognition of excellence and competitiveness. While this strategy offers a broader perspective, it inherently excludes individuals who, despite their expertise, have not received such awards. It is important to acknowledge the limitations of this validation approach. By focusing solely on award recipients, there is a risk of overlooking highly influential individuals who have not been formally recognized. This limitation underscores the need for a comprehensive validation strategy that encompasses a diverse range of experts, regardless of their award status. In comparison with other validation methods that may be more institution-centric, this approach offers a broader scope but still falls short in capturing the entirety of expertise within a given field. Nevertheless, its emphasis on globally recognized achievements adds a valuable dimension to the evaluation process.

4 Result and discussion

In this section, we explained the experimental results derived from our study, utilizing the mathematics dataset.

4.1 Simple ranking

We conducted an analysis to explore how various indices impact the ranking of awardees at the top of the list. To accomplish this, we investigated the presence of awardees within the top 100 positions of the ranked list for each parameter. After collecting the necessary data, we assessed the rankings of authors for each index individually and recorded the positions of the award winners within the ranked lists. Additionally, we calculated the number of award winners who secured a place within the top 100 of each list. The findings presented in Fig. 3 indicate that among the award winners who ranked within the top 100 authors, the top 5 parameters specifically, K-index, Cite/paper, platinum h-index, h2 lower index, and A-index demonstrated the highest occurrence, ranging from 64 to 78 percent. In contrast, total publications, gF index, and gm index exhibited the weakest performance, with only 2% of awardees.

4.2 Ranking with deep learning

In this section, we ranked the parameters based on their importance score. This importance score extracted by using MLP classifier with modified recursive elimination technique (explained in methodology section). The importance score of a parameter represents the overall impact of this parameter on model performance during classification of awardees and non-awardees. Figure 4 represents the ranking of all parameters based on importance score. In Figure 4, it is evident that among the parameter ranking based on importance score, the top 5 parameters—namely normalized hi index, maxprod, gf index, F index, and hi index—showed

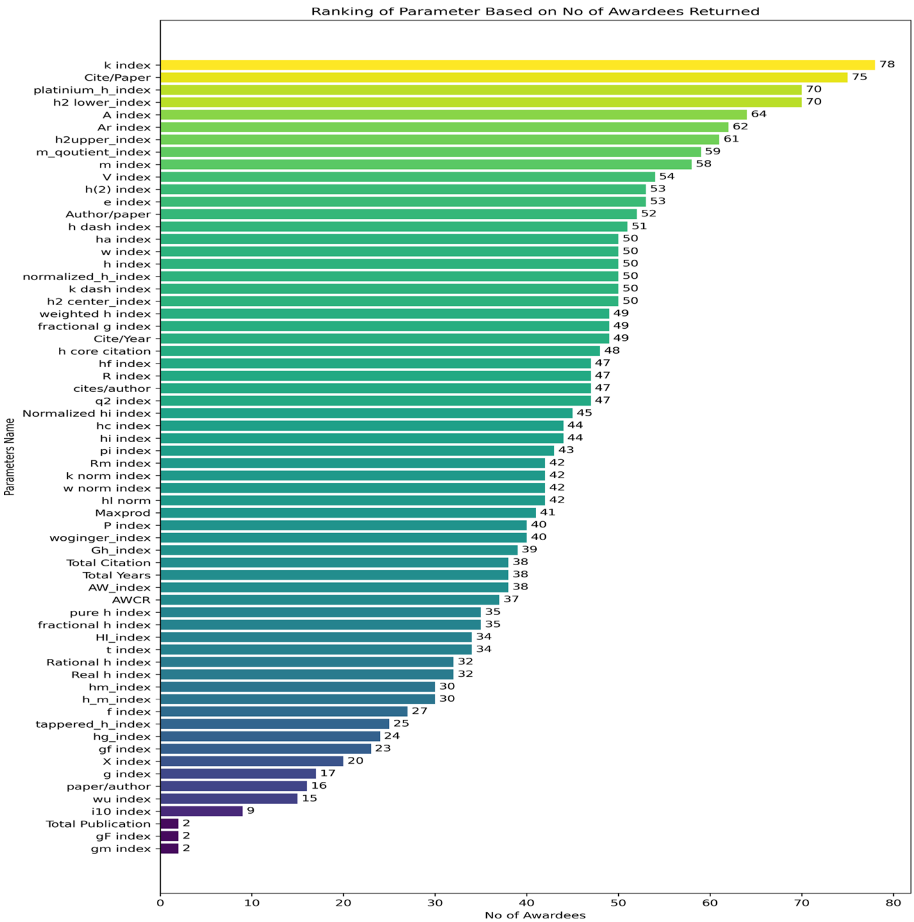

Fig. 3 Simple ranking of parameter based on no of returned awardees in to top 100 records

the highest importance score, ranging from 0.17 to 0.22. Conversely, k index exhibited the lowest impact on model performance, with importance score 0.01.

4.3 Calculating disjointness ratio

In this section, we began by selecting the top-performing parameters and generating all possible combinations. Subsequently, for each of these combinations, we compiled author ranking lists for both parameters. From these lists, we then extracted the top 100 records and identified occurrences of awardees. Within the occurrences of awardees, we computed the degree of disjointness, representing the percentage of unique awardees returned by each pair of parameters. Table 3 provides a comprehensive overview of the disjointness results for all parameter pairs. In Table 3, the first two columns present the parameter pair, the third column indicates the number of unique awardees returned by the first parameter, followed by the number of unique awardees returned by the second parameter. Furthermore, the fifth column displays the common awardees present in both parameter lists, while the sixth and seventh

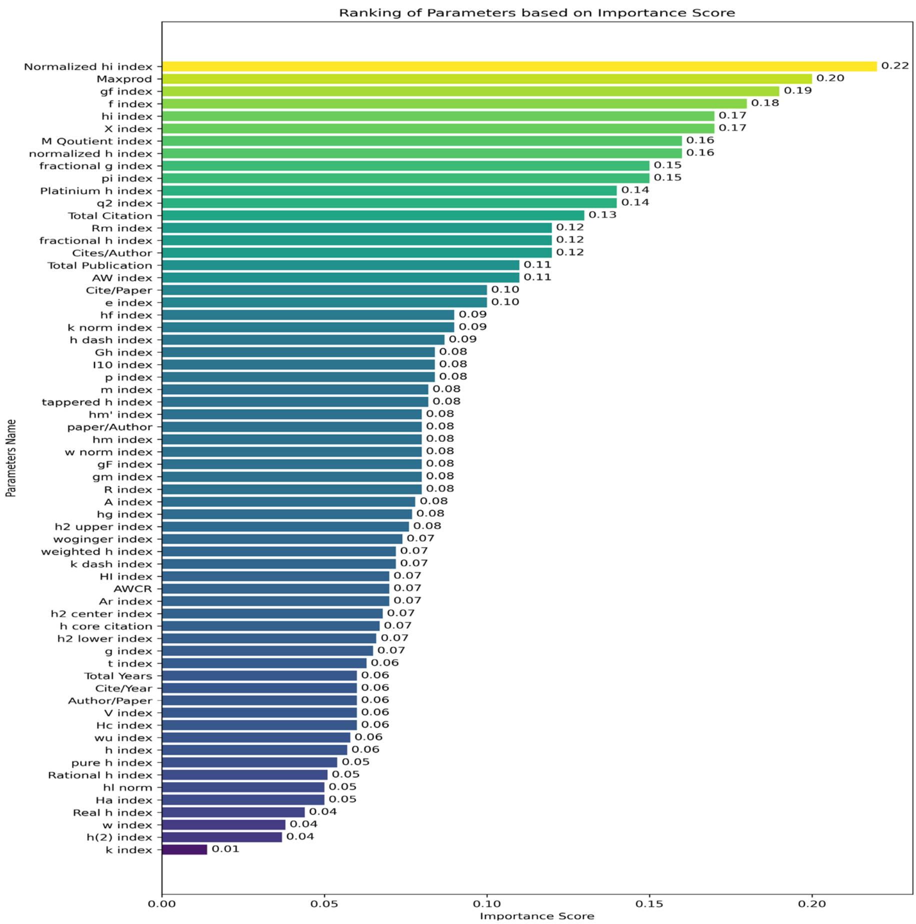

Fig. 4 Ranking of Parameter based on importance Score

columns show the total number of awardees and the total number of unique awardees. The last column represents the disjointness ratio between the parameter pairs. Analyzing Table 1 reveals that the "gf index" and "k index" exhibit the highest disjointness ratio of 0.96, while the combination of "normalized hi index" and "hi index" yields the lowest disjointness ratio, which is zero.

4.4 Statistical analysis and proposed index

This section involved combining the top highest disjoint pair (gf index and k index) across eight different statistical models. This process aims to explore and identify the top statistical model which retain the maximum characteristic of this pair. By employing this comprehensive approach, we can gain valuable insights into the interplay between this pair and statistical models, thereby enhancing our understanding of their combined effects. For this we have to calculate the values of eight different statistical methods against this pair. After obtaining the values for each statistical method, we sorted the list of the values generated by each

Table 3 Disjointness result

1st parameter	2nd parameter	1st unique	2nd unique	Both common	Total returned	Total unique	Disjointness ratio
gf index	K index	21	76	2	99	97	0.96
Cite/Paper	gf index	70	38	3	111	108	0.95
gf index	H2 lower index	18	66	5	89	84	0.89
F index	K index	20	71	7	98	91	0.87
Cite/Paper	F index	73	29	8	110	102	0.86
Cite/Paper	hi index	69	42	12	123	111	0.82
Cite/Paper	Normalized hi index	70	41	12	123	111	0.82
gf index	Platinum index	14	66	9	89	80	0.82
F index	H2 lower index	17	61	10	88	78	0.8
H2 lower index	hi index	59	32	12	103	91	0.79
H2 lower index	Normalized hi index	59	32	12	103	91	0.79
hi index	K index	30	64	14	108	94	0.77
K index	Normalized hi index	64	30	14	108	94	0.77
gf index	hi index	16	36	8	60	52	0.76
gf index	Normalized hi index	15	36	8	59	51	0.76
F index	A index	15	52	12	79	67	0.74
gf index	Maxprod	13	31	10	54	44	0.69
H2 lower index	K index	48	55	23	126	103	0.69
F index	gf index	19	15	8	42	34	0.68
K index	Maxprod	57	20	21	98	77	0.65
A index	H2 lower index	40	47	24	111	87	0.64
A index	hi index	44	24	20	88	68	0.63
A index	Normalized hi index	44	20	24	88	64	0.63
hi index	Maxprod	28	25	16	69	53	0.62

Table 3 continued

1st parameter	2nd parameter	1st unique	2nd unique	Both common	Total returned	Total unique	Disjointness ratio
Maxprod	Normalized hi index	25	28	16	69	53	0.62
F index	Platinum index	7	55	20	82	62	0.61
A index	gf index	47	6	17	70	53	0.61
H2 lower index	Maxprod	49	19	22	90	68	0.61
hi index	Platinum index	21	52	23	96	73	0.61
Normalized hi index	Platinum index	21	52	23	96	73	0.61
Cite/Paper	Maxprod	64	14	27	105	78	0.59
F index	hi index	12	29	15	56	41	0.58
F index	Normalized hi index	12	29	15	56	41	0.58
K index	Platinum index	40	37	38	115	77	0.5
A index	K index	28	42	36	106	70	0.49
H2 lower index	Platinum index	34	38	37	109	72	0.49
Cite/Paper	Platinum index	48	32	43	123	80	0.48
Cite/Paper	A index	48	21	43	112	69	0.45
F index	Maxprod	8	22	19	49	30	0.44
Maxprod	Platinum index	8	42	33	83	50	0.43
A index	Platinum index	24	35	40	99	59	0.42
A index	Maxprod	32	9	32	73	41	0.39
Cite/Paper	K index	38	25	53	116	63	0.37
Cite/Paper	H2 lower index	38	18	53	109	56	0.35
hi index	Normalized h index	44	44	44	132	88	0

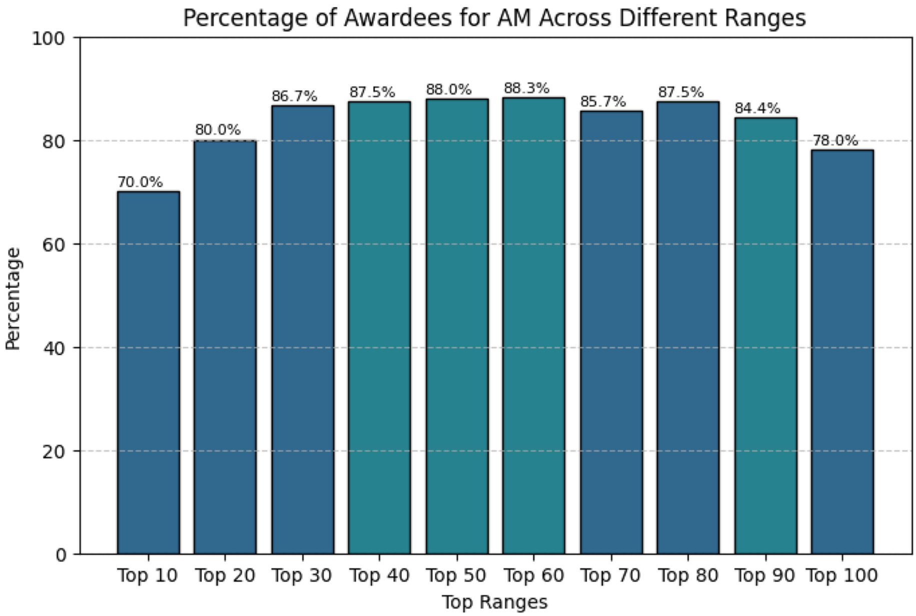

Fig. 5 Arithmetic mean

method. After sorting, to further investigate the performance of these statistical methods, we conducted an analysis of the top 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 records. This analysis involved examining the number of awardees returned by each statistical method within these subsets of data. By observing the outcomes, we can evaluate the effectiveness of statistical methods for identifying awardees. The model-wise result is presented below:

4.4.1 Arithmetic mean (AM)

Figure 5 presents the results of arithmetic mean (AM) model. From this figure, it is evident that the model demonstrates a high degree of accuracy, with the percentage of awardees identified remaining consistently above 80% across all ranges. The model performs best within the top 30 records, peaking at an 88% success rate. Notably, there is a slight decrease in the model's performance as the range of records expands, with the lowest percentage of 78% occurring at the top 100 records mark. This pattern suggests that while the AM model is robust in identifying awardees within a smaller, more concentrated dataset, its predictive power marginally diminishes as the dataset broadens. Nonetheless, the model's performance remains impressively high throughout, indicating its potential utility in award prediction scenarios where a high degree of accuracy is required within the top echelons of candidates.

4.4.2 Harmonic mean (HM)

Figure 6 presents the results of harmonic mean (AM) model. The HM model's predictive accuracy exhibits a positive trend as the range of records increases, with the percentage of awardees identified starting at 20% for the top 10 records and gradually rising to a peak of 55.7% for the top 70 records. Beyond this point, the model's performance plateaus, maintaining a success rate slightly above 55% for the top 80, top 90, and top 100 records. This plateau

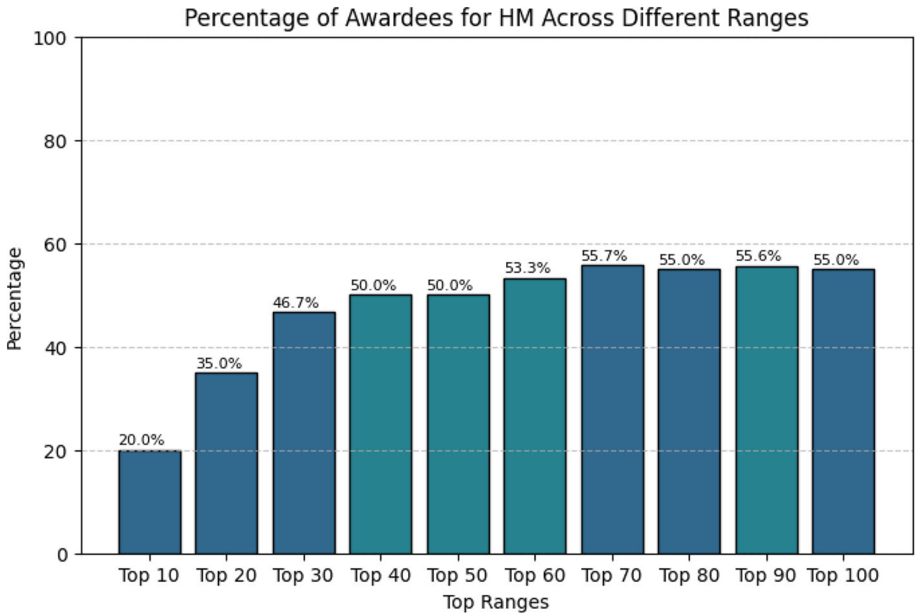

Fig. 6 Harmonic mean

suggests that while the HM model benefits from a larger dataset, its predictive capability stabilizes and does not significantly improve after a certain threshold. The overall pattern indicates that the HM model is more effective when analyzing a broader dataset but has an upper limit to its predictive accuracy, which is important to consider when applying this model in practical scenarios.

4.4.3 Contra-harmonic mean (CHM)

Figure 7 presents the results of contra-harmonic mean (CHM) model. The CHM model demonstrates a high level of accuracy, with a notable initial success rate of 70% for the top 10 records. This rate of correct predictions increases as the record range expands, reaching an optimal performance at the top 70 records with an 87.1% success rate. Beyond this peak, the model experiences a slight decrease in effectiveness, with the percentage of awardees identified dropping to 81% at the top 100 records mark. The trend suggests that the CHM model is exceptionally effective within a moderate range of records, but its predictive accuracy shows a marginal decline when the dataset is expanded to its fullest extent. This pattern is crucial for researchers and practitioners who consider employing the CHM model for award prediction, as it indicates the model's strengths within certain data range limits.

4.4.4 Geometric mean (GM)

Figure 8 presents the results of geometric mean (GM) model. The GM model starts with a high success rate of 90% for the top 10 records, which is consistent up to the top 30 records. The percentage slightly decreases to 88% for the top 40 records and shows a gradual decline as the range increases, with the lowest percentage being 84% for the top 100 records. This

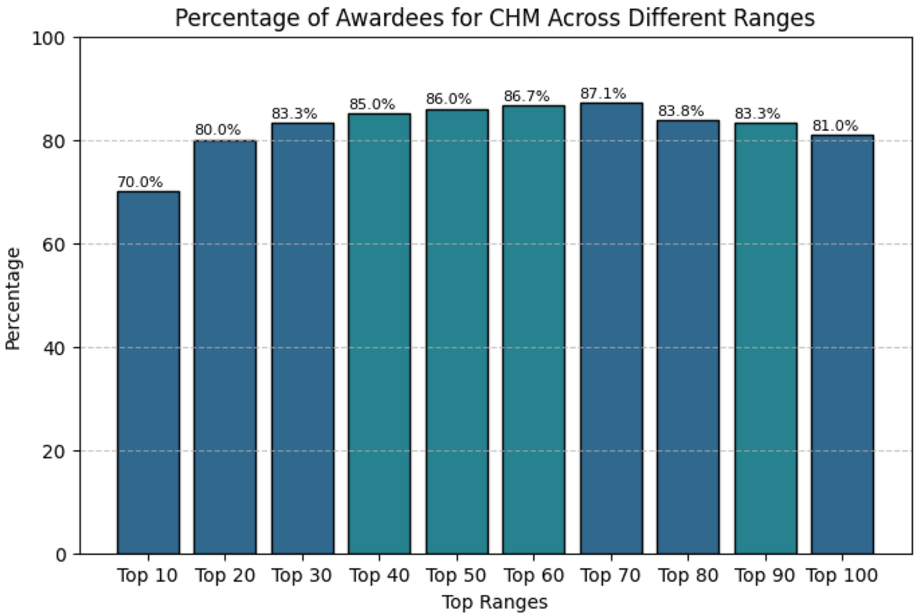

Fig. 7 Contra Harmonic Mean

trend indicates that the GM model is highly effective in identifying awardees within a smaller dataset, and while its performance slightly diminishes with larger datasets, it still maintains a relatively high prediction accuracy throughout the ranges.

4.4.5 Logarithmic mean (LM)

Figure 9 presents the results of logarithmic mean (LM) model. The LM model exhibits a commendable predictive accuracy, initiating at a 70% success rate for the top 10 records and reaching a peak accuracy of 90% for the top 50 records. The model maintains a high level of performance across the board, although a slight decline in predictive accuracy is observed as the dataset expands, with the success rate dipping to 80% for the top 100 records. This trend indicates that the LM model is particularly potent in the middle ranges of the dataset, with a slight reduction in effectiveness at the dataset's extremities. The data suggest that the LM model could be a valuable tool for predicting awardees, especially when applied to a focused range of top-performing candidates.

4.4.6 Root mean square (RMS)

Figure 10 presents the results of logarithmic mean (LM) model. The RMS model shows a strong start with 70% accuracy for the top 10 records, and this accuracy improves as the range increases, peaking at 88% for the top 50 records. The model maintains high performance in the 80% range for most record categories, with a slight decrease to 78% for the top 100 records. This pattern suggests that the RMS model is quite robust in its predictive ability, especially within the mid-range of the dataset, and only shows a minor drop in performance when applied to the largest range of records.

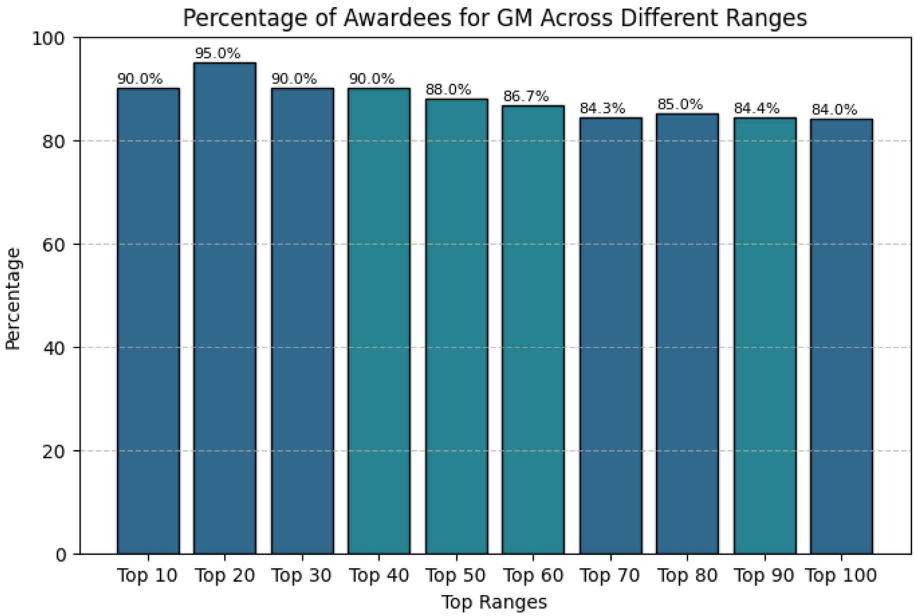

Fig. 8 Geometric mean

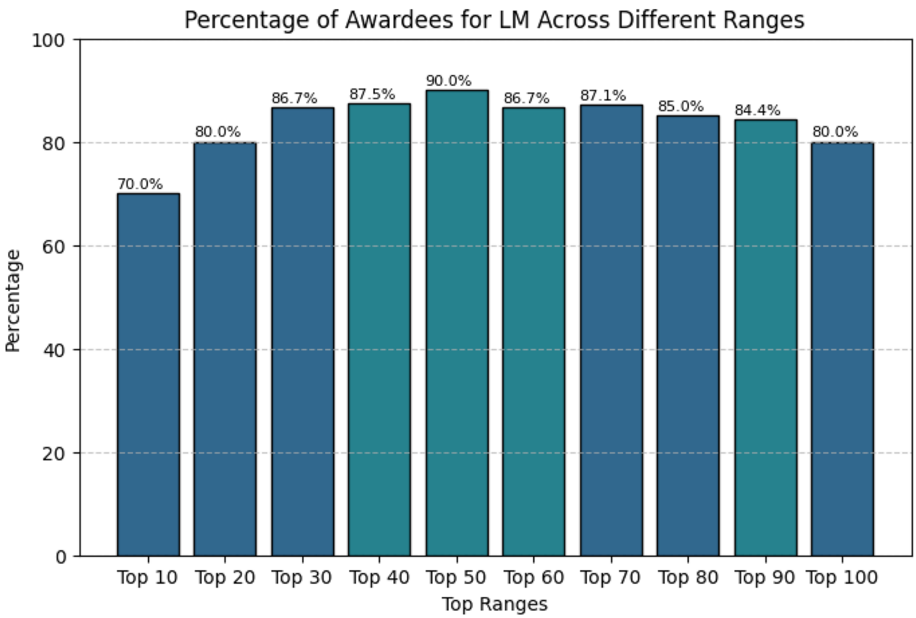

Fig. 9 Logarithmic mean

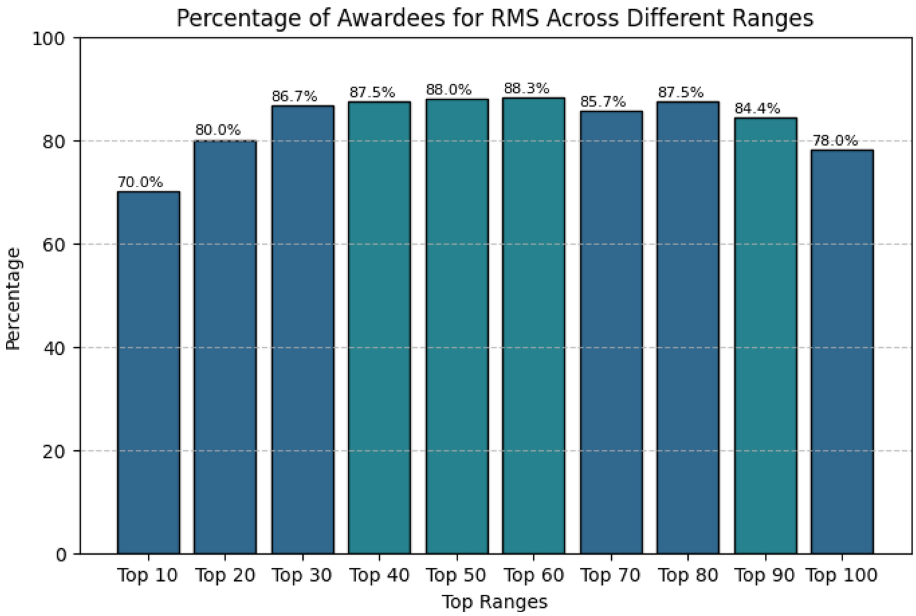

Fig. 10 Root mean square

4.4.7 Trigonometric mean (TM)

Figure 11 presents the results of trigonometric mean (TM) model. The result reveals that the TM model achieves its highest predictive accuracy of 46% for the top 50 records, indicating a strong performance within a mid-range dataset. While the model's effectiveness fluctuates across different ranges, it maintains a relatively stable success rate around 40% for larger datasets. This suggests that the TM model is reasonably effective across various ranges, with its optimal performance observed within a more concentrated set of top-performing records.

4.5 Statistical model comparison

Figure 12 presents the comparisons of statistical model. From the graph, we can observe the following:

1. The AM (blue line) and HM (green line) models show very similar performance across all ranges, maintaining a high percentage that slightly decreases as the range increases. They start off close to 90% and end just below this mark at the Top 100 range, indicating robust performance across the board.
2. The CHM (brown line) and GM (red line) models also exhibit similar trends, starting at around 80% and slightly declining as the range increases, ending just above 80%. These models are consistently less effective than the AM and HM models but still show strong performance.
3. The LM (purple line) model starts at a lower percentage around 70% but maintains a relatively stable performance across all ranges, ending just below 70%. This model is less effective than the AM, HM, CHM, and GM models.

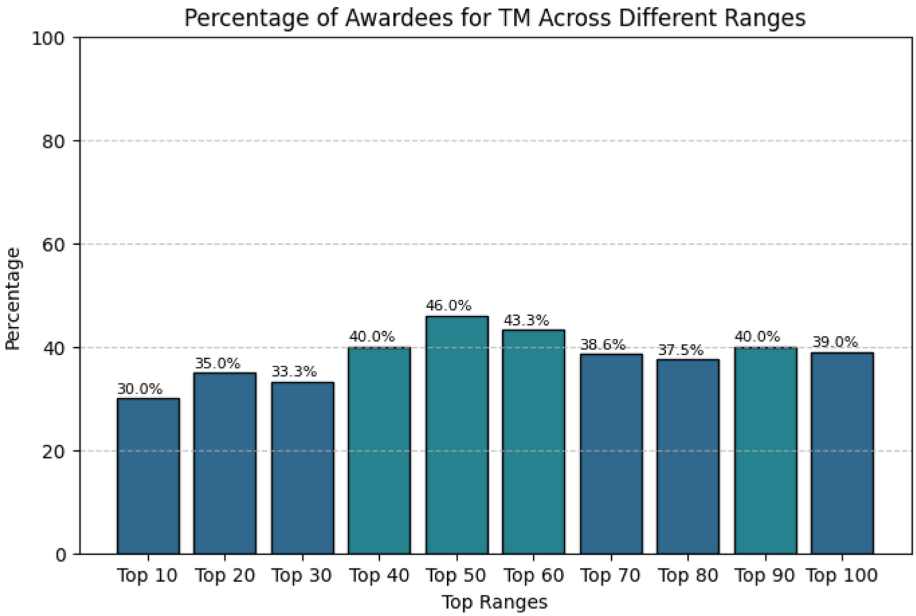

Fig. 11 Trigonometric mean

4. The RMS (orange line) model shows a different trend, starting at around 40% and increasing performance as the range increases, reaching just above 50%. This suggests that the RMS model may be more suited to broader data ranges rather than more specific segments.
5. The TM (pink line) model starts the lowest, around 30%, and shows a slight increase before plateauing and then decreasing again, ending around 40%. This model is the least effective across all ranges when compared to the other models.

Moreover, in comparative analysis of the statistical models, Figure 12 distinctly illustrates the varying average impacts of each model. The geometric mean (GM) emerges as the most influential with an average impact score of 87.74, suggesting its robustness in handling multiplicative data relationships. The arithmetic mean (AM), logarithmic mean (LM), and root mean square (RMS) are closely aligned, each with an average impact of approximately 83.6, indicating their consistent reliability across diverse datasets. The contraharmonic mean (CHM) follows closely with an impact of 82.62, reinforcing its utility in averaging rates. In stark contrast, the harmonic mean (HM) and trigonometric mean (TM) register significantly lower impacts of 47.63 and 38.27, respectively. This contrast underscores the nuanced applicability of these models, with HM and TM potentially offering more specialized insights in datasets characterized by outliers or non-normal distributions. The visual representation through Fig. 13 not only aids in understanding the relative performance of each model but also guides the selection of the appropriate statistical model based on the dataset characteristics and the research question at hand. Based on the analysis, it is evident that the GM model surpasses all statistical models and is suitable for combining the top disjoint pairs while preserving their maximum properties. However, directly inputting the gf and k indexes into the geometric mean formula assumes that both indexes contribute equally to predicting whether the awardee receives an award or not. This assumption is contradictory given that the two

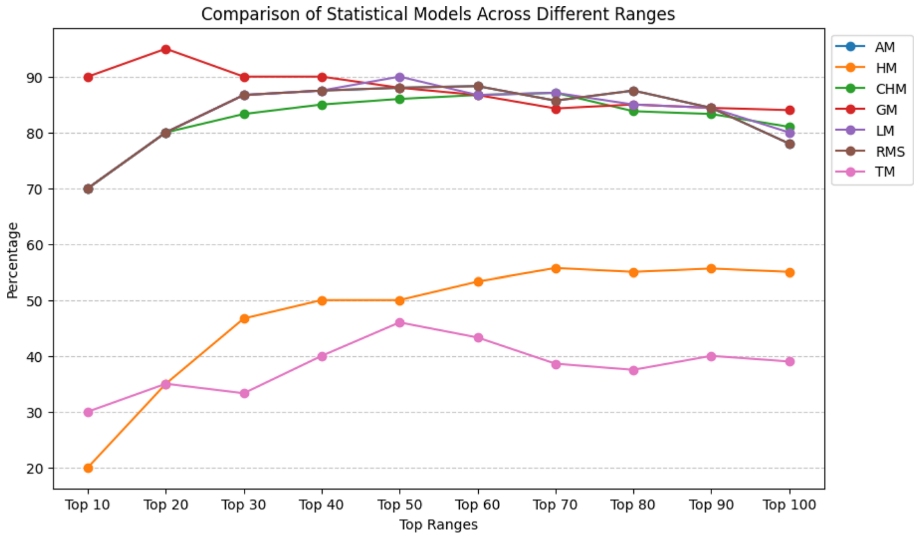

Fig. 12 Comparison between different statistical model

indexes resulted in different numbers, indicating that they affect the prediction differently. To address this, we need to compute the weight associated with this top-performing pair, signifying the importance of an index compared to others in the formula. For weight calculation, we have already ranked all indices using two approaches, where an index's performance in these rankings is considered its weight. This means that an index's position relative to the total number of indices (64 in this case) determines its weight. Equation 3 provided below indicates that as an index moves down the ranking, its weight decreases. Essentially, the index competes with the other 64 indices, and its position in the ranking determines its weight.

$$I_W = Avg\left(\frac{\frac{SR_p}{\frac{1}{T_i}}}{T_i}, \frac{\frac{DR_p}{\frac{1}{T_i}}}{T_i}\right) \tag{3}$$

In the above formula, the I_W represents index weight, the SR_p represents the position of index in simple ranking, DR_p represents the position of index in deep learning ranking and T_i represents the total no of indices.
We utilize the above equation to compute the weights for the K index and the gf index. These weights pertain to the field of mathematics; should the field change, these weights will be recalculated accordingly. Subsequently, we combine both indices using the geometric mean, introducing a novel index known as the GK index. The GK index is defined as the geometric mean of the gf index and the K index, as illustrated in the equation below.

$$GKindex = \sqrt{0.16 * gfindex * 0.25 * Kindex} \tag{4}$$

5 Conclusion

In this paper, we address the evolving challenges in ranking scholars within the scientific community by critically evaluating an extensive array of parameters and proposing a novel

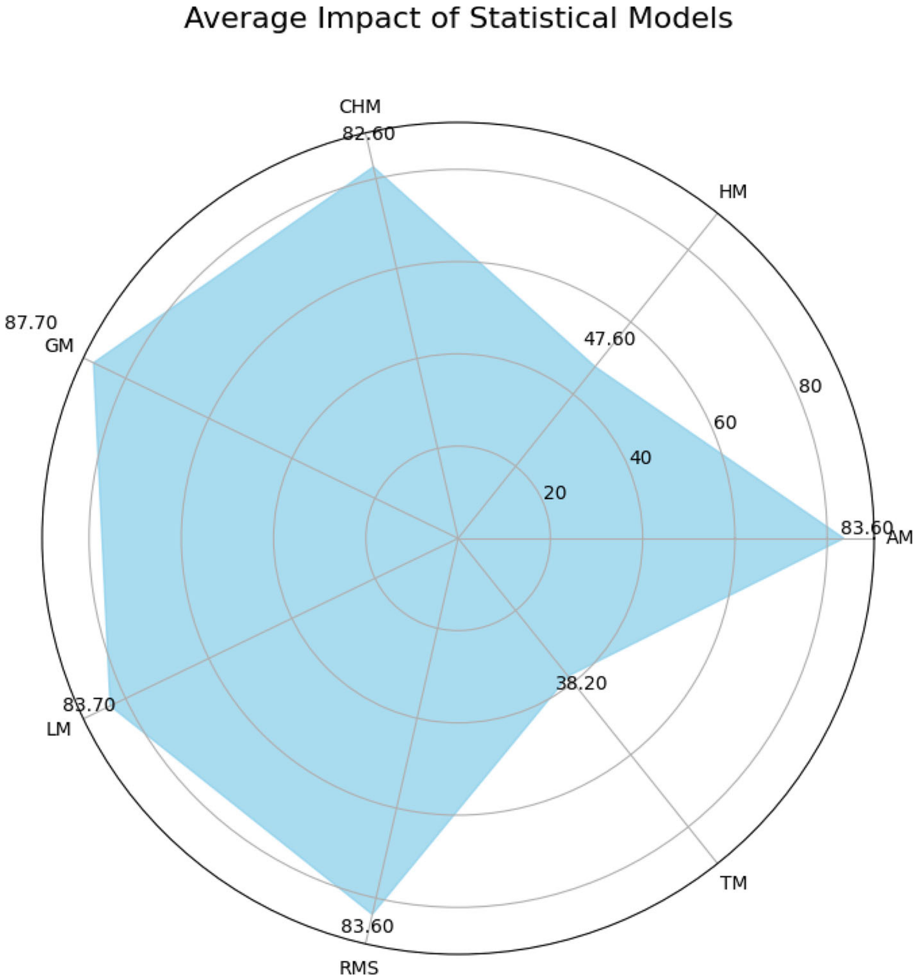

Fig. 13 Average impact of statistical considering all ranges

index. Despite the existing literature proposing over 70 parameters for author ranking, encompassing both quantitative and qualitative aspects, a consensus within the scientific community remains elusive regarding a single parameter for ranking researchers. To tackle this issue, we adopt a comprehensive approach, considering sixty-four parameters and employing both traditional ranking based on calculated index values and advanced deep learning techniques. To mitigate bias, instead of selecting the top parameter, we identify the top five parameters that consistently demonstrate success in identifying awardees. For experimental purposes, we utilize a mathematics domain dataset. Additionally, we calculate the disjointness ratio between the top parameter and all possible combinations, selecting the top two parameters with the highest disjointness ratio. Subsequently, the evaluation of these parameters through seven statistical models allows us to identify the most influential model, providing a robust foundation for combining the top disjoint pair. We also integrate weight factors with the index formula to represent the importance of one index over the other.

In conclusion, we introduce the GK index, a novel metric devised by amalgamating the top disjoint pair, consisting of the gf index and the K index. This integration is facilitated through the utilization of the most influential statistical model, specifically the geometric mean. As the scientific community undergoes continual evolution, our study not only contributes to the current dialogue but also establishes a valuable framework for prospective discussions and advancements in the assessment of scholarly impact. There are several limitations to this study. Firstly, the proposed novel index is only applicable to the mathematics domain. As the field changes, the proposed index may also need to change, requiring recalculation of its weight factor. Secondly, our validation method focuses on researchers who have received awards, potentially excluding experts who have not received awards.

6 Future work

In our future endeavors, we aspire to expand the scope of our research in two key dimensions. Firstly, we are incorporating additional published indices into our list of metrics, such as the Kaptay K index [46], the h α-index [47], the ψ-index [48], Q-factor [49], EM index [50], and Year-based h-type index [51]. Secondly, we are procuring datasets from a wide array of domains, encompassing fields such as Civil Engineering, Neuroscience, and Computer Science, among others. This expansion will enable a thorough evaluation of the sixty-four parameters, alongside the newly proposed parameter, across diverse disciplines.

Appendix A

Table 4 Indices calculation formulas

Name of index	Calculation
Total Publication [8]	Total No of Publication of a researcher
Total Citation [9]	Total No of Citation of a researchers
Total Years [40]	Total Number of years since the researchers first publication
Cites/Year [40]	Total citations count / Total Number of years since first paper
Cites/Paper [40]	Total citations/total papers
Author/Paper [40]	Calculate the sum of authors contributing to the publications associated with the specified author, then divide this total by the number of papers
Cites/author [40]	Compute the citations per publication by dividing the citations for each paper by the number of authors, and then sum these adjusted citation values. This sum represents the single-authored equivalent number of citations for the specified author
Papers/author [40]	Calculate the single-authored equivalent number of papers for the specified author by dividing each publication by the number of authors and summing the fractional author counts

Acknowledgements This research is not funded by any organization. Muhammad Tanvir Afzal gave the idea and supervise the research, Ghulam Mustafa did the experiments, interpreted the results and wrote the paper, Abid Rauf review the paper and supervise the research.

Table 4 continued

Name of index	Calculation
H-index [12]	h = max (numbers of publication with >= h citations count)
G-index [23]	A collection of papers is assigned a g-index, denoted as 'g,' where 'g' is the highest rank for which the top g papers collectively amass at least g^2 citations
Hg-index [52]	$hg - index = \sqrt{h * g}$ where h represents h index and g represents g index
A-index [53]	$A - index = \frac{1}{h}\sum_{p=1}^{h} cit_p$ where A is the A-index of the scholar, h represents the h-index and cit_p is the citation count of the p^{th} article
R-index [19]	$R - index = \sqrt{\sum_{p=1}^{h} Cit_p}$ where h represents the h-index and cit_p is the citation count of the p^{th} article
P-index [8]	$P - index = (\frac{C^2}{p})^{\frac{1}{3}}$ The p-index strikes the best balance between total number of citations (C) and the mean citation rate (C/P)
Q^2-index [10]	$Q^2 = \sqrt{h * m}$ where h represents h index and m represents m index
K-index [54]	$K - index = \frac{\frac{C}{P}}{\frac{C(h-tail)}{C(h-core)}}$ where C represents total citation, p represents of p^{th} article, h-tail represents h-tail article citation and h-core represents h-core citation
E-index [55]	$e - index = \sum_{p=1}^{h} Cit_p - h^2$ where Cit_p represents citation of p^{th} article and h represents h-index
f-index [18]	$f - index = \binom{max}{f}\frac{1}{\frac{1}{f}\sum_{p=1}^{f}\frac{1}{Cit_p}} \geq f$ where cit_p is the citation count of p^{th} article. The f-index never goes beyond the total number of publications
T-index [18]	$T - index = \binom{max}{t} exp\left[\frac{1}{t}\sum_{k=1}^{t} ln(Cit_k)\right] \geq t$ where cit_k is the citation count of k^{th} article
Tapered h-index [56]	$H_{T(j)} = \begin{cases} \frac{n_j}{2j-1} & n_j \leq j \\ h_{T(j)} = \frac{j}{2j-1} + \sum_{i=j+1}^{n_j}\frac{1}{2j-1} & n_j > j \end{cases}$
Wu index [57]	The w-index of an author is determined by calculating it as follows: if at least w of their articles have accumulated 10w citations each, while the remaining publications have garnered fewer than 10(w+1) citations each
Weighted h-index [58]	$R_w(k) = \frac{\sum_{p=1}^{k} \mathrm{Cit}_p}{h}$ where h is the h-index and Cit_p is the citation count of the p^{th} article. Then, the weighted h-index is defined as follows $h_w = \sqrt{\sum_{k=1}^{R} \mathrm{Cit}_k}$ where Cit_k is the citation count of the k^{th} article and R is the largest rank among all publications such that the k^{th} weighted rank $\leq \mathrm{Cit}_k$

Table 4 continued

Name of index	Calculation
h(2)-index [39]	The h(2)-index is defined as the highest integer 'h(2)' for which the scholar's top h(2) most cited articles each have received at least $(h(2))^2$ citations
Woeginger index [59]	$w = \binom{max}{w}(Cit_p \geq w - p + 1) \quad for\ all\ \ p \leq w$ where $cit_p$is the citation count of p^{th} article and w is the maximum number of publications
Rm-index [60]	$R_m = \sqrt{\sum_{k=1}^{h} \mathrm{Cit}_k^{\frac{1}{2}}}$ where cit_k is the citation count of k^{th} article
m-index [61]	The m-index is computed as the median number of citations received by the h-core articles
X-index [62]	$x = \sqrt{\binom{max}{k} kCit_k}$ where cit_k is the citation count of k^{th} article
h2 upper-index [63]	$h^2upper = \frac{\sum_{k=1}^{h}(Cit_k - h)}{\sum_{k=1}^{m} Cit_k} * 100 = \frac{e^2}{\sum_{k=1}^{m} Cit_k} * 100$ where h is the h-index, cit_k is the citation count of the k^{th} article, $e^2$is the excess citation and m is the total number of articles
h2 center-index [63]	$h^2center = \frac{h*h}{\sum_{k=1}^{m} Cit_k} * 100$ where h is the h-index and cit_k is the citation count of the k^{th} article
h2 lower-index [63]	$h^2lower = \frac{\sum_{k=h+1}^{m}(Cit_k - h)}{\sum_{k=1}^{m} Cit_k} * 100$ where h is the h-index and cit_k is the citation count of the k^{th} article
h̀-index (h dash) [64]	$h^{'} = Rh = \frac{eh}{t}$ where R represents the head-tail ratio of e and t-index
Rational h-index [65]	$h_{rat} = h + 1 - \frac{k}{2h+1}$ Where h is the h-index and k is the number of citations required to reach h+1 h-index value
I10 index [66]	This is a straightforward and direct indexing measure that entails counting the total number of papers published by a journal, each of which has received at least 10 citations
Normalized h-index [67]	$normalized\, h - index = \frac{h}{pub_{count}}$ where h represents h index and pubcount represents total publication
$\Pi\, index$ [68]	$\Pi\, index = 0.01C(P\Pi)$ The $\Pi - index$ is equal to the 100th of the number of citations, $C(P\pi)$ to the top square root $(P\pi)$ of the total papers(P)
Gh-index[69]	$Gh^a = \sum_{p=1}^{m} sing(Cit(\mathrm{pub}_b^a) - GH) \quad where\ sing(x) = 1, x \geq 0\, and\ 0, x \leq 0$ where 'm' represents the total number of publications by scholar A, and GH is the h-index of the scholar. This index is also challenging to compute compared to the h-index

Table 4 continued

Name of index	Calculation
W index [70]	The w-index is defined by the parameter 'w,' signifying the number of a scholar's top articles that have accrued at least 10w citations each. While the w-index can be a valuable measure for gauging the impact of a scholar, it may impose a penalty on young scholars who have recently commenced their work or those who have not yet published a sufficient number of papers
Maxprod [71]	The maximum value of i * c_i can be determined by analyzing the publication rank of an author, where c_i represents the number of citations for the i^{th} most frequently cited paper among all the citations
H core citation [40]	The h-core citation index considers only publications that have been cited at least h times, disregarding those that have not reached this threshold
K dash index [72]	$k' = \frac{Cit_{all} - Pub_{Count}}{Cit_T - Cit_H}$ where cit_{all} represents total citation, pub_{count} represents total publication, cit_t represents total citation of h tail article and cit_h represents total citation of h core article
M-Quotient [73]	$M - Quotient = \frac{h-index}{y}$ where y represents no of the year the first publication
Hc - index [74]	hc-index $= \alpha \cdot \frac{C(i)}{Y(\text{now}) - Y(i) + 1}$ where $Y(\text{now})$ represents the current year, $Y(i)$ represents the publication year, and $C(i)$ represents the citation count of paper i hc-index $= \frac{C(i)}{1}, \frac{C(i)}{2}, \frac{C(i)}{3}, \ldots, \frac{C(i)}{n}$
Aw-index [75]	Aw-index $= \sqrt{\sum_{j=1}^{h} \frac{\text{Cit}_j}{a_j}}$ where Cit_j represents the citation count of the j-th article, and a_j represents the j-th article out of m total articles
Ar index [19]	$Ar - index = \sqrt{\sum_{j=1}^{h} \frac{Cit_j}{a_j}}$ where Cit_j citation of the article and a_j represents a^{th} article, h represents total h core article
AWCR [76]	This parameter adjusts the citation count based on the length of time that has passed since each publication
v-index [77]	$V = \frac{h}{P(y_{this} - y_o)}$ where h is the h-index, y_{this} is the current year and y_0 is the year of first publication
Platinum H-index [78]	$Platinum - h = \frac{H}{CL} * \frac{cit_{all}}{Pub_{count}}$ where H is the h-index, CL is the career length, Cit_{all} is the total citation count and pub_{count} is the publication count
Ha index [79]	The ha-index within a dataset is the highest count of papers in the dataset that have garnered at least ha citations per year on average
HI Index [80]	$h_i = \frac{h}{Avg_a}$ where h represents h index and Avg_a represents average no of authors in article
hf index [81]	$\frac{Yh_f}{\phi(Yh_f)} \geq h_f$

Table 4 continued

Name of index	Calculation
	where Y(i) represents citation count and $\phi(i)$ represents average no of authors in article
gf index [81]	$gf = \sum_{i=1}^{gf} \frac{Y_i}{\phi_i} \geq g^2 f$
	where Y(i) represents citation count and $\phi(i)$ represents average no of authors in article
gF index [81]	$gF = \left(\sum_{i=1}^{k} \frac{1}{\phi(i)}\right)^2 \leq \sum_{i=1}^{k} y_i$
	where Y(i) represents citation count and $\phi(i)$ represents average no of authors in article
Normalized Hi index [82]	normalized hi index $= \frac{h}{pub_{count}}$
	where h represents the h-index and pub_{count} is the total number of articles..
Hm index [83]	$r_{eff}(r) = \sum_{r'=1}^{r} \frac{1}{a(r')} thenc(r(h_m)) \geq h_m \geq c(r(h_m)+1)$
k norm index [84]	$k - norm = h - norm + (1 - (h - \frac{norm^2}{\sum_{j=1}^{h-norm} citnorm_i})), \quad \forall h - norm > 1\, and\, k - norm = 0,\ if\ h - norm = 0$
w norm index [84]	$w - norm = h - norm + (1 - (h - \frac{norm^2}{totalcit-norm})), \quad \forall h - norm > 0\, and\, w - norm = \frac{totalcit-norm}{1+totalcit-norm},\ if\ h - norm = 0$
gm index [85]	$g_m \leq C_{eff}(g_m)\ where\ C_{eff}(r_{eff})\ \ and\ S_{eff)}(r_{eff}) = \sum_{\acute{r}=1}^{r(r_{eff})} \frac{1}{a(\acute{r})} c(\acute{r})$
pure h index [86]	$h_p(A) = \frac{h}{\sqrt{E(author)}}$
	where h represents h index and E average no of author
fractional h-index [87]	$h_f = max(k \leq \frac{cit(k)}{author(k)})$
	where Cit_k represents citation of article and author(K) represents no of author in a specific article
fractional g-index [87]	$g_f = max(\sum_{k=1}^{p} \frac{cit_k}{Author(k)} \geq p^2)$
hi norm index [88]	The hl-norm is a modified version of the h-index that normalizes citations, taking into account the number of authors per paper
Real h index [89]	$h_r = \frac{(h+1)cit_h - h.cit_h + 1}{1 - cit_{h+1} + cit_h}$
	where h is the h-index and cit_h is the citation count of the h^{th} article

References

1. Mustafa G, Usman M, Yu L, Afzal MT, Sulaiman M, Shahid A (2021) Multi-label classification of research articles using word2vec and identification of similarity threshold. Sci Rep 11(1):21900. https://doi.org/10.1038/s41598-021-01460-7
2. Xia W, Li T, Li C (2023) A review of scientific impact prediction: tasks, features and methods. Scientometrics 128(1):543–585. https://doi.org/10.1007/s11192-022-04547-8
3. Jiang X, Sun X, Zhuge H (2013) Graph-based algorithms for ranking researchers: not all swans are white! Scientometrics 96:743–759. https://doi.org/10.1007/s11192-012-0943-y
4. Ahmed B, Li W, Mustafa G, Afzal MT, Alharthi SZ, Akhunzada A (2023) Evaluating the effectiveness of author-count based metrics in measuring scientific contributions. IEEE Access. https://doi.org/10.1109/ACCESS.2023.3309416
5. Raheel M, Ayaz S, Afzal MT (2018) Evaluation of h-index, its variants and extensions based on publication age & citation intensity in civil engineering. Scientometrics 114:1107–1127. https://doi.org/10.1007/s11192-017-2633-2

6. Bihari A, Tripathi S, Deepak A (2023) A review on h-index and its alternative indices. J Inf Sci 49(3):624–665. https://doi.org/10.1177/01655515211014478
7. Harzing A-WK, Wal R (2008) Google scholar as a new source for citation analysis. Ethics Sci Environ Polit 8(1):61–73. https://doi.org/10.3354/esep00076
8. Prathap G (2010) The 100 most prolific economists using the p-index. Scientometrics 84(1):167–172. https://doi.org/10.1007/s11192-009-0068-0
9. Liu J-X, Yin M-M, Gao Y-L, Shang J, Zheng C-H (2022) Msf-lrr: multi-similarity information fusion through low-rank representation to predict disease-associated microbes. IEEE/ACM Trans Comput Biol Bioinf 20(1):534–543. https://doi.org/10.1109/TCBB.2022.3146176
10. Cabrerizo FJ, Alonso S, Herrera-Viedma E, Herrera F (2010) q2-index: quantitative and qualitative evaluation based on the number and impact of papers in the hirsch core. J Informet 4(1):23–28. https://doi.org/10.1016/j.joi.2009.06.005
11. Mustafa G, Rauf A, Al-Shamayleh AS, Ahmed B, Alrawagfeh W, Afzal MT, Akhunzada A (2023) Exploring the significance of publication-age-based parameters for evaluating researcher impact. IEEE Access. https://doi.org/10.1109/ACCESS.2023.3304013
12. Hirsch JE (2005) An index to quantify an individual's scientific research output. Proc Natl Acad Sci 102(46):16569–16572. https://doi.org/10.1073/pnas.050765510
13. Dienes KR (2015) Completing h. J Informet 9(2):385–397. https://doi.org/10.1016/j.joi.2015.01.003
14. Mustafa G, Rauf A, Ahmed B, Afzal MT, Akhunzada A, Alharthi SZ (2023) Comprehensive evaluation of publication and citation metrics for quantifying scholarly influence. IEEE Access 11:65759–65774. https://doi.org/10.1109/ACCESS.2023.3290917
15. Ayaz S, Afzal MT (2016) Identification of conversion factor for completing-h index for the field of mathematics. Scientometrics 109(3):1511–1524. https://doi.org/10.1007/s11192-016-2122-z
16. Ain Q-U, Riaz H, Afzal MT (2019) Evaluation of h-index and its citation intensity based variants in the field of mathematics. Scientometrics 119:187–211. https://doi.org/10.1007/s11192-019-03009-y
17. Chen Z, Yao J, Xiao G, Wang S (2021) Efficient and differentiable low-rank matrix completion with back propagation. IEEE Trans Multimed. https://doi.org/10.1109/TMM.2021.3124087
18. Tol R (2009) The h-index and its alternatives: An application to the 100 most prolific economists. Scientometrics 80(2):317–324. https://doi.org/10.1007/s11192-008-2079-7
19. Jin B, Liang L, Rousseau R, Egghe L (2007) The r-and ar-indices: complementing the h-index. Chin Sci Bull 52(6):855–863. https://doi.org/10.1007/s11434-007-0145-9
20. Zhang C-T (2009) The e-index, complementing the h-index for excess citations. PLoS ONE 4(5):5429. https://doi.org/10.1371/journal.pone.0005429
21. Aziz NA, Rozing MP (2013) Profit (p)-index: the degree to which authors profit from co-authors. PLoS ONE 8(4):59814. https://doi.org/10.1371/journal.pone.0059814
22. Burrell QL (2007) On the h-index, the size of the hirsch core and jin's a-index. J Inf 1(2):170–177. https://doi.org/10.1016/j.joi.2007.01.003
23. Egghe L (2006) Theory and practise of the g-index. Scientometrics 69(1):131–152. https://doi.org/10.1007/s11192-006-0144-7
24. Lopez J, Susarla SM, Swanson EW, Calotta N, Lifchez SD (2015) The association of the h-index and academic rank among full-time academic hand surgeons affiliated with fellowship programs. J Hand Surg 40(7):1434–1441. https://doi.org/10.1016/j.jhsa.2015.03.026
25. Bihari A, Tripathi S, Deepak A (2023) A review on h-index and its alternative indices. J Inf Sci 49(3):624–665. https://doi.org/10.1177/01655515211014478
26. Khan NR, Thompson CJ, Taylor DR, Gabrick KS, Choudhri AF, Boop FR, Klimo P Jr (2013) Part ii: should the h-index be modified? an analysis of the m-quotient, contemporary h-index, authorship value, and impact factor. World Neurosurg 80(6):766–774. https://doi.org/10.1016/j.wneu.2013.07.011
27. Katsaros D, Akritidis L, Bozanis P (2009) The f index: quantifying the impact of coterminal citations on scientists' ranking. J Am Soc Inform Sci Technol 60(5):1051–1056. https://doi.org/10.1002/asi.21040
28. Cameron DHL, Aleman-Meza B, Decker S, Arpinar IB (2007) Semef: a taxonomy-based discovery of experts, expertise and collaboration networks. PhD thesis, University of Georgia. https://doi.org/10.1177/01655515211014478
29. Ye F, Rousseau R (2010) Probing the h-core: an investigation of the tail-core ratio for rank distributions. Scientometrics 84(2):431–439. https://doi.org/10.1007/s11192-009-0099-6
30. Kosmulski M et al (2006) A new hirsch-type index saves time and works equally well as the original h-index. ISSI Newslett 2(3):4–6. https://doi.org/10.1177/01655515211014478
31. Van Raan AF (2006) Comparison of the hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. scientometrics 67:491–502. https://doi.org/10.1556/Scient.67.2006.3.10

32. Jin B, Liang L, Rousseau R, Egghe L (2007) The r-and ar-indices: complementing the h-index. Chin Sci Bull 52(6):855–863. https://doi.org/10.1007/s11434-007-0145-9
33. Mane K, Shrawankar U (2023) An improved indexing technique for tribal art retrieval system. In: 2023 IEEE international students' conference on electrical, electronics and computer science (SCEECS), pp. 1–5. https://doi.org/10.1109/SCEECS57921.2023.10061818
34. Xiao S, Yan J, Li C, Jin B, Wang X, Yang X, Chu SM, Zha H (2016) On modeling and predicting individual paper citation count over time. In: Ijcai, pp. 2676–2682. https://doi.org/10.1109/SCEECS57921.2023.10061818
35. Wu X (2021) W-index: a weighted index for evaluating research impact. Open J Appl Sci 11:149–156. https://doi.org/10.4236/ojapps.2021.111010
36. Ayaz S, Afzal MT (2016) Identification of conversion factor for completing-h index for the field of mathematics. Scientometrics 109(3):1511–1524. https://doi.org/10.1007/s11192-016-2122-z
37. Ameer M, Afzal MT (2019) Evaluation of h-index and its qualitative and quantitative variants in neuroscience. Scientometrics 121(2):653–673. https://doi.org/10.1007/s11192-019-03209-6
38. Salman M, Ahmed MM, Afzal MT (2021) Assessment of author ranking indices based on multi-authorship. Scientometrics 126(5):4153–4172. https://doi.org/10.1007/s11192-021-03906-1
39. Usman M, Mustafa G, Afzal MT (2021) Ranking of author assessment parameters using logistic regression. Scientometrics 126(1):335–353. https://doi.org/10.1007/s11192-020-03769-y
40. Alshdadi AA, Usman M, Alassafi MO, Afzal MT, AlGhamdi R (2023) Formulation of rules for the scientific community using deep learning. Scientometrics 128(3):1825–1852. https://doi.org/10.1007/s11192-023-04633-5
41. Ghani R, Qayyum F, Afzal MT, Maurer H (2019) Comprehensive evaluation of h-index and its extensions in the domain of mathematics. Scientometrics 118:809–822. https://doi.org/10.1007/s11192-019-03007-0
42. Harzing A-WK, Wal R (2008) Google scholar as a new source for citation analysis. Eth Sci Environ Polit 8(1):61–73. https://doi.org/10.3354/esep00076
43. Yin Y, Jang-Jaccard J, Xu W, Singh A, Zhu J, Sabrina F, Kwak J (2023) Igrf-rfe: a hybrid feature selection method for mlp-based network intrusion detection on unsw-nb15 dataset. J Big Data 10(1):1–26. https://doi.org/10.1186/s40537-023-00694-8
44. Awad M, Fraihat S (2023) Recursive feature elimination with cross-validation with decision tree: feature selection method for machine learning-based intrusion detection systems. J Sens Actuator Netw 12(5):67. https://doi.org/10.3390/jsan12050067
45. Kilincer IF, Ertam F, Sengur A, Tan R-S, Acharya UR (2023) Automated detection of cybersecurity attacks in healthcare systems with recursive feature elimination and multilayer perceptron optimization. Biocybern Biomed Eng 43(1):30–41. https://doi.org/10.1016/j.bbe.2022.11.005
46. Kaptay G (2020) The k-index is introduced to replace the h-index to evaluate better the scientific excellence of individuals. Heliyon. https://doi.org/10.1016/j.heliyon.2020.e04415
47. Hirsch JE (2019) hα: an index to quantify an individual's scientific leadership. Scientometrics 118(2):673–686. https://doi.org/10.1007/s11192-018-2994-1
48. Lathabai HH (2020) ψ-index: a new overall productivity index for actors of science and technology. J Informet 14(4):101096. https://doi.org/10.1016/j.joi.2020.101096
49. Sinatra R, Wang D, Deville P, Song C, Barabási A-L (2016) Quantifying the evolution of individual scientific impact. Science 354(6312):5239. https://doi.org/10.1126/science.aaf5239
50. Bihari A, Tripathi S, Deepak A (2023) A review on h-index and its alternative indices. J Inf Sci 49(3):624–665
51. Mahbuba D, Rousseau R (2013) Year-based h-type indicators. Scientometrics 96:785–797. https://doi.org/10.1007/s11192-012-0934-z
52. Alonso S, Cabrerizo F, Herrera-Viedma E, Herrera F (2010) hg-index: A new index to characterize the scientific output of researchers based on the h-and g-indices. Scientometrics 82(2):391–400. https://doi.org/10.1007/s11192-009-0047-5
53. Jin B (2006) H-index: an evaluation indicator proposed by scientist. Sci Focus 1(1):8–9. https://doi.org/10.1209/0295-5075/78/30002
54. Ye F, Rousseau R (2010) Probing the h-core: an investigation of the tail-core ratio for rank distributions. Scientometrics 84(2):431–439. https://doi.org/10.1007/s11192-009-0099-6
55. Zhang C-T (2009) The e-index, complementing the h-index for excess citations. PLoS ONE 4(5):5429. https://doi.org/10.1371/journal.pone.0005429
56. Anderson TR, Hankin RK, Killworth PD (2008) Beyond the durfee square: enhancing the h-index to score total publication output. Scientometrics 76:577–588. https://doi.org/10.1007/s11192-007-2071-2
57. Wu Q (2010) The w-index: a measure to assess scientific impact by focusing on widely cited papers. J Am Soc Inform Sci Technol 61(3):609–614. https://doi.org/10.1002/asi.21276

58. Egghe L, Rousseau R (2008) An h-index weighted by citation impact. Inf Process Manag 44(2):770–780. https://doi.org/10.1016/j.ipm.2007.05.003
59. Woeginger GJ (2008) An axiomatic characterization of the hirsch-index. Math Soc Sci 56(2):224–232. https://doi.org/10.1016/j.mathsocsci.2008.03.001
60. Panaretos J, Malesios C (2009) Assessing scientific research performance and impact with single indices. Scientometrics 81:635–670. https://doi.org/10.1007/s11192-008-2174-9
61. Bornmann L, Mutz R, Daniel H-D (2008) 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. J Am Soc Inform Sci Technol 59(5):830–837. https://doi.org/10.1002/asi.20806
62. Fenner T, Harris M, Levene M, Bar-Ilan J (2018) A novel bibliometric index with a simple geometric interpretation. PLoS ONE 13(7):0200098. https://doi.org/10.1371/journal.pone.0200098
63. Bornmann L, Mutz R, Daniel H-D (2010) The h index research output measurement: two approaches to enhance its accuracy. J Informet 4(3):407–414
64. Zhang C-T (2013) The h'-index, effectively improving the h-index based on the citation distribution. PLoS ONE 8(4):59912. https://doi.org/10.1371/journal.pone.0059912
65. Ruane F, Tol R (2008) Rational (successive) h-indices: an application to economics in the republic of Ireland. Scientometrics 75(2):395–405. https://doi.org/10.1007/s11192-007-1869-7
66. Scholar G. Measuring your research impact: i10-index. Google Scholar https://guides.library.cornell.edu/impact
67. Sidiropoulos A, Katsaros D, Manolopoulos Y (2007) Generalized hirsch h-index for disclosing latent facts in citation networks. Scientometrics 72:253–280. https://doi.org/10.1007/s11192-007-1722-z
68. Vinkler P (2009) The π-index: a new indicator for assessing scientific impact. J Inf Sci 35(5):602–612. https://doi.org/10.1177/0165551509103601
69. Xu F, Liu W, Mingers J (2015) New journal classification methods based on the global h-index. Inf Process Manag 51(2):50–61. https://doi.org/10.1016/j.ipm.2014.10.011
70. Wu, Q.: The w-index: A significant improvement of the h-index. arXiv preprint arXiv:0805.4650 (2008)
71. Kosmulski M (2007) Maxprod-a new index for assessment of the scientific output of an individual, and a comparison with the h-index. Cyberm Int J Scientomet Inf Bibliomet 11:5. https://doi.org/10.4103/1008-682X.171582
72. Chen D-Z, Huang M-H, Fred YY (2013) A probe into dynamic measures for h-core and h-tail. J Informet 7(1):129–137. https://doi.org/10.1016/j.joi.2012.10.002
73. Burrell QL (2007) Hirsch's h-index: a stochastic model. J Informet 1(1):16–25. https://doi.org/10.1016/j.joi.2006.07.001
74. Sidiropoulos A, Katsaros D, Manolopoulos Y (2007) Generalized hirsch h-index for disclosing latent facts in citation networks. Scientometrics 72:253–280. https://doi.org/10.1007/s11192-007-1722-z
75. Agarwal A, Durairajanayagam D, Tatagari S, Esteves SC, Harlev A, Henkel R, Roychoudhury S, Homa S, Puchalt NG, Ramasamy R et al (2016) Bibliometrics: tracking research impact by selecting the appropriate metrics. Asian J Androl 18(2):296. https://doi.org/10.4103/1008-682X.171582
76. Cucchetti A, Mazzotti F, Pellegrini S, Cescon M, Maroni L, Ercolani G, Pinna AD (2013) The use of the hirsch index in benchmarking hepatic surgery research. Am J f Surg 206(4):560–566. https://doi.org/10.1016/j.amjsurg.2013.01.037
77. Vaidya JS (2005) V-index: a fairer index to quantify an individual's research output capacity. BMJ. https://doi.org/10.1209/0295-5075/78/30002
78. Smith DR (2015) "Platinum H": refining the H-index to more realistically assess career trajectory and scientific publications. https://doi.org/10.1080/19338244.2015.1016833
79. Hagen NT (2010) Harmonic publication and citation counting: sharing authorship credit equitably-not equally, geometrically or arithmetically. Scientometrics 84(3):785–793. https://doi.org/10.1007/s11192-009-0129-4
80. Batista PD, Campiteli MG, Kinouchi O (2006) Is it possible to compare researchers with different scientific interests? Scientometrics 68(1):179–189. https://doi.org/10.1007/s11192-006-0090-4
81. Egghe L (2008) Mathematical theory of the h-and g-index in case of fractional counting of authorship. J Am Soc Inform Sci Technol 59(10):1608–1616. https://doi.org/10.1002/asi.20845
82. Wohlin C (2009) A new index for the citation curve of researchers. Scientometrics 81(2):521–533. https://doi.org/10.1007/s11192-008-2155-z
83. Schreiber M (2008) A modification of the h-index: the hm-index accounts for multi-authored manuscripts. J Informet 2(3):211–216. https://doi.org/10.1016/j.joi.2008.05.001
84. Anania G, Caruso A (2013) Two simple new bibliometric indexes to better evaluate research in disciplines where publications typically receive less citations. Scientometrics 96:617–631. https://doi.org/10.1007/s11192-013-0951-6

85. Schreiber M (2009) Fractionalized counting of publications for the g-index. J Am Soc Inform Sci Technol 60(10):2145–2150. https://doi.org/10.1002/asi.21119
86. Wan J-K, Hua P-H, Rousseau R (2007) The pure h-index: calculating an author'sh-index by taking co-authors into account. COLLNET J Sci Inf Manag 1(2):1–5. https://doi.org/10.1080/09737766.2007.10700824
87. Egghe L (2008) Mathematical theory of the h-and g-index in case of fractional counting of authorship. J Am Soc Inform Sci Technol 59(10):1608–1616. https://doi.org/10.1002/asi.20845
88. Alonso S, Cabrerizo FJ, Herrera-Viedma E, Herrera F (2009) h-index: A review focused in its variants, computation and standardization for different scientific fields. J Informet 3(4):273–289. https://doi.org/10.1016/j.joi.2009.04.001
89. Guns R, Rousseau R (2009) Real and rational variants of the h-index and the g-index. J Informet 3(1):64–71. https://doi.org/10.1016/j.joi.2008.11.004

Ghulam Mustafa is currently doing his PhD in Computer Science from University of Engineering and Technology, Taxila, Pakistan. He has received his MS in Computer Science degree with distinction from Capital University of Science and Technology, Islamabad, and secured Gold medal due to his excellent academic performance in his entire degree duration. Moreover, in his BS in Software Engineering degree COMSATS, Abbottabad has awarded two Gold Medals on his 1st position in Abbottabad campus, as well as got 1st position in all seven campuses of CUI in 2017. He has been associated with academia, and industry for the last 6 years, and currently, he is serving as Senior Lecturer in Faculty of Computing at Shifa Tameer-e-Millat University, Park Road, Islamabad. Previously, he has served 5 years in CS department of Capital University of Science and technology, Islamabad, as Associate and Junior Lecturer. Furthermore, he has been served A&F Solution Software House as a web frontend designer and backend developer. Moreover, with academia he is also an active freelancer from the last 4 years, doing projects in different languages such as, python, java, and C++.

Abid Rauf received the MSc degree in Statistics from the Quaid-i-Azam University, Islamabad, Pakistan, in 2000, the MS degree in information security from Sichuan University, Chengdu, China, in 2007, and PhD degree in 2021 from National University of Sciences and Technology (NUST), Islamabad, Pakistan. He is currently working as a lecturer in department of Computer Science, University of Engineering and Technology, Taxila, Pakistan. He has expertise in the area of information security and privacy. His research interests include cryptographic protocols and algorithms, security in Internet of Things (IoT), security in wireless sensor networks and privacy, machine learning, and privacy and covert communications.

Dr. Tanvir Afzal received the PhD degree with highest distinction in Computer Science from the Graz University of Technology, Austria, and secured Gold medal in his MSc Computer Science from Quaid-i-Azam University, Islamabad, Pakistan. He has been associated with academia and industry at various levels for the last 20 years, and currently, he is serving as Director and Professor, Shifa School of Computing and Director Campus at one of the largest campus of Shifa Tameer-e-Millat University at park Road Islamabad. Previously, he has served Namal Institute Mianwali as Director and Professor. He has served Capital University of Science and Technology, Islamabad, as Professor, Associate Professor, and Assistant Professor of CS. Furthermore, he has served NESCOM, COMSATS University Islamabad, and JinTech Islamabad. He is also serving as editor-in-chief for reputed impact factor journal known as: Journal of Universal Computer Science. Dr. Afzal authored more than 125 research papers including over 50 published papers 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 123+, with citations over 1,300. He played pivotal role in making collaborations between MAJU-JUCS, MAJU-IICM, and TUG-UNIMAS. He served as PhD symposium chair, session chair, finance chair, committee member, and editor of several IEEE, ACM, Springer, Elsevier international conferences and journals. Dr. Afzal conducted more than 100 curricular, cocurricular, and extracurricular activities in the last 5 years including seminars, workshops, national competitions (ExcITeCup) and invited international and national speakers from Google, Oracle, IICM, IFIS, SEGA Europe, etc. Under his supervision, more than 70 post-grad students (MS and PhD) have defended their research theses successfully and a number of PhD and MS students are pursuing their research with him. He served as master trainer and a 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 master trainer for data science trainings in leading industrial organizations. He is the recipient of multiple international research fundings.