*COLLNET Journal of Scientometrics and Information Management*
ISSN : 0973 – 7766 (Print) 2168 – 930X (Online)
Vol. 19(1) June 2025, pp. 91–122
DOI : 10.47974/CJSIM-2024-11007

# Formulating rules to identify key researchers in computer science : A quantitative approach

**Ghulam Mustafa**
**Muhammad Tanvir Afzal**

**In the vast expanse of scientific literature, determining which researchers have made the most significant contributions to their field is essential. However, the scientific community lacks a standardized set of criteria for identifying impactful researchers. Existing platforms, like Google Scholar, Semantic Scholar & Web of Science, provide useful metrics like total publications, citation counts, and h-index, but no universally accepted framework exists to comprehensively evaluate a researcher's true impact. This study proposes a novel framework for identifying impactful researchers within the computer science domain. The framework integrates 64 distinct quantitative parameters across categories such as citation-based, publication-based, author-count-based, and age-weighted metrics to evaluate researchers' contributions. Unlike traditional methods that focus on individual indicators, our framework provides a comprehensive evaluation model that considers multiple factors simultaneously. The parameters were selected based on their importance scores, determined by their ability to classify awardees and non-awardees. Experiments were performed on a balanced dataset comprising 600 awardees and 600 non-awardees, yielding classification accuracies ranging from 57% to 78% in recognizing influential researchers. The top-ranked features from each parameter category effectively promoted 50% to 55% of awardees into the top 100 researcher rankings. These findings offer a robust tool for academic institutions and organizations seeking to identify impactful researchers. The proposed framework enables more objective evaluation and provides actionable insights for individual researchers striving for recognition and influence in their field.**

***Keywords:*** *Researcher evaluation metrics, Parameter prioritization, Multi-layer perceptron (MLP), H-index variants, Computer science domain datasets.*

**Ghulam Mustafa** *
**Muhammad Tanvir Afzal** †
Department of Computer Science
Shifa Tameer-e-Millat University (STMU)
Islamabad 44000
Pakistan

***E-mail***
* ghulam.mustafa.ssc@stmu.edu.pk
† dean.foc@stmu.edu.pk

* (Corresponding Author)

## 1. Introduction

Each day, millions of researchers contribute to the expanding body of scientific knowledge [1, 2]. However, the true significance and quality of these contributions often remain unclear, as the impact of research is typically recognized only after a delay. This delay hampers timely acknowledgment of researchers deserving of accolades, limiting their early influence on the academic community. Consequently, assessing research impact has become a critical and widely debated topic, leading to the development of various evaluation frameworks. Despite these efforts, no universally accepted model for recognizing exceptional scientific contributions exists. Several evalu- ation techniques have been proposed, ranging from peer assessments to metric-based analyses [3][4][5]. While expert review lends credibility, it is often time-consuming and resource-heavy. Conversely, quantitative metrics such as citation counts and publication volume offer efficiency but can be misleading. Scholars may inflate metrics through self-citation or by publishing in lower-tier venues [6][7][8].

A significant shift occurred with Hirsch's introduction of the h-index, designed to account for both publication quantity and citation quality [9, 10]. While widely adopted, the h-index has received criticism. Dienes noted that additional citations to existing h-core publications do not necessarily lead to an increase in the h-index [11, 12]. Moreover, articles with comparable citation counts outside the h-core are often overlooked [13, 14]. These limita- tions have prompted the development of alternative indices [15]. Over 70 such metrics have emerged, including the A-index [16], AR-index [17], M-Quotient [18], and k-index [19], but evaluating these indices across hypothetical or narrowly defined datasets has hindered the identification of the most practically relevant metrics [20, 21].

To address these shortcomings, various empirical studies have assessed the performance of newly proposed indices. As an example, Dienes [11] explored the use of the h-index, g-index, and complementary h-index to evaluate research impact within the field of mathematics[11, 22], while De, Nayeem, and Pal [23] focused on civil engineering [23], and Schreiber et al. [24] extended their analysis to neuroscience [24]. Other evaluations, such as those by Ain et al. [20], Ghani et al. [25], and Moreira, Calado, and Martins [26], examined multiple indices across various fields to identify reliable measures of scholarly influence. Recent works by Kumar [27] and Jose and Franklin [28].

Ahmed et al. [29] have built on these studies, incorporating factors like publication age and author count into their evaluations, while Usman et al. [6] and Ahmed et al. [5] employed machine learning models like logistic regression and random forests to rank parameters and refine evaluation strategies [5].

Despite this extensive body of work, there is still no comprehensive set of decision rules that can guide researchers or institutions in prioritizing relevant evaluation parameters. These rules are critical for identifying which metrics best reflect research excellence and for providing tangible goals for aspiring researchers. This gap highlights the need for clear, actionable frameworks to facilitate researcher evaluation and award nominations [30–32].

While numerous prior studies have investigated the connection between bibliometric indicators particularly various forms of the h index and award recognition [27, 28, 31, 33], the majority have been constrained by their reliance on a limited and often homogeneous set of parameters. For instance, the work by Alshdadi et al. [7] represents a notable attempt to define interpretable rules for researcher impact; however, their framework was confined to a small number of features, limiting the generalizability and depth of their findings. In contrast, the present study adopts a more comprehensive and fine-grained approach by evaluating a significantly larger set of indicators. Specifically, we incorporate a total of 64 bibliometric metrics, carefully grouped into four distinct cat- egories: (1) primitive metrics, (2) publication and citation based indicators, (3) author count sensitive metrics, and

(4) age weighted indices. This expanded parameter space enables a more robust and nuanced analysis of researcher impact. To determine the significance of each metric, we utilized a Multi-Layer Perceptron (MLP) classifier in conjunction with Recursive Feature Elimination (RFE), enabling a data-driven approach to feature ranking. From each metric category, the five most influential parameters were selected. Subsequently, we used a Decision Tree classifier to generate interpretable decision rules based on these top-ranked features. This methodological pipeline not only facilitates transparent rule extraction but also supports practical applications in both individual researcher evaluation and institutional policy making. All experiments were conducted on a balanced dataset comprising 600 award-winning and 600 non-awarded researchers in the computer science domain, ensuring fairness and statistical reliability in performance assessment.

In this study, we seek to explore and answer the following key research questions that guide our investigation:

- Which specific quantitative bibliometric indicators have the strongest influence on the likelihood of a researcher receiving a prestigious award?
- How do combinations of parameters (rather than individual thresholds) determine award eligibility?

The insights gained from this work can serve as foundational guidelines for establishing formal criteria in sci- entific recognition processes. Additionally, they provide early-career researchers with a structured path to enhance their scholarly impact.

The remainder of this paper is organized to provide a logical and comprehensive flow of the research. Section 2 offers an in-depth review of the existing literature, highlighting key contributions, limitations, and research gaps in the study of bibliometric indicators and researcher recognition. Section 3 introduces the proposed methodology, detailing the framework for selecting, ranking, and evaluating bibliometric parameters, as well as the techniques used to derive interpretable decision rules. Section 4 presents the experimental results, including performance evaluations, insights from the ranked features, and implications of the derived rules. Section 5 draws key con- clusions from the study, emphasizing its contributions to the field of researcher evaluation and impact analysis.

Finally, Section 6 outlines potential directions for future work, including the integration of dynamic metrics, domain adaptation, and extensions to interdisciplinary datasets.

## 2. Literature

In today's competitive research environment, the demand for objective, transparent, and standardized mechanisms for assessing academic performance is stronger than ever. Traditional indicators such as total publication count, citation frequency, and the widely adopted h-index are commonly used to evaluate scholarly productivity and influence. Despite their popularity, these metrics are often insufficient for capturing the complexity of academic contributions. For instance, academic awards and recognitions frequently reflect subjective assessments, and yet these are often anchored in quantitative metrics that may not be appropriate across disciplines [34–38]. One major criticism is that raw publication counts can overestimate impact, especially if the publications appear in venues with limited visibility or rigor [39, 40]. Similarly, citation-based measures can be misleading due to self-citations, reciprocal citation networks, or even citations that arise from negative critiques [41–43].

Among the existing indicators, the h-index has emerged as a benchmark for evaluating a balance between productivity and impact. Its appeal lies in its simplicity and the ease with which it can be computed. However, scholars have noted several inherent limitations in its design. For example, once a paper is included in the h-core, additional citations to it do not influence the index's value [44]. Additionally, the h-index often disadvantages early-career researchers due to their limited citation history, while potentially overstating the influence of senior or inactive researchers whose past publications continue to attract citations[45–47]. In response to these short- comings, several supplementary and alternative indicators—such as the A-index, k-index, and f-index—have been proposed to provide a more balanced and refined evaluation of scholarly impact [48, 49].

A growing number of studies have sought to evaluate the effectiveness of these indicators across diverse academic domains. For example, Ayaz et al. [15] analyzed award datasets from mathematical societies and concluded that the h-index remained superior to other available alternatives. Building upon this, Raheel et al. [31] explored additional index variants and reinforced the h-index's predictive capacity. In the neuro- science domain, Ameer et al. [50] identified the hg-index and R-index as strong discriminators of research excellence. Similarly, Ain et al. [51] conducted an evaluation in the field of mathematics and found signif- icant correlations between certain metrics and award recognition. However, one limitation of these earlier studies is their reliance on outdated award datasets, some of which predate the introduction of several modern bibliomet- ric indices, thereby reducing the relevance of the findings. Addressing this issue, Usman et al. [2] shifted focus to the civil engineering field and analyzed more temporally aligned award and non-award data. Neverthe- less, the limited sample size in their dataset restricted the ability to generalize their findings, indicating a broader issue with data accessibility and scale in metric validation studies.

Additional contributions have focused on developing rule-based models to enhance the interpretability of researcher evaluations. Alshdadi et al. [7], for instance, proposed a framework based on rule induction, although it relied on a relatively narrow set of metrics and lacked multi-dimensional assessment capabilities. In a more domain-focused effort, Mustafa et al. [27, 28] conducted two parallel investigations that delved into distinct sets of features. One study emphasized publication and citation-based metrics and concluded that the normalized h-index demonstrated the highest predictive validity. The other study concentrated on metrics related to publication age, identifying the Ar-index as a particularly effective indicator. Notably, both studies were limited to the mathematics domain, which raises questions about the generalizability of their findings to other fields.

The evolution of researcher assessment tools reveals a gradual shift from basic, count-based indicators toward more sophisticated and context-aware metrics. While early evaluations relied heavily on total publications and raw citation numbers, more recent research emphasizes refined indices—especially variations of the h-index—that aim to capture additional dimensions of academic influence. However, these newer indices are not without their own constraints, particularly in terms of domain sensitivity and the effects of temporal bias. Furthermore, inconsistent methodologies, differing sample populations, and a lack of unified benchmarks across studies make it difficult to draw broad conclusions.

Taken together, the literature underscores a critical need for robust, multi-dimensional evaluation frameworks that can accommodate diverse research profiles and institutional contexts. Such frameworks should integrate multiple metric types ranging from primitive and citation-based indicators to age- and author-count-adjusted in- dices—to construct a holistic picture of academic impact. These systems not only enable fairer recognition of deserving researchers but also have the potential to inform funding decisions, promotion criteria, and policy formulation in higher education and research institutions. The current study addresses this gap by proposing a structured, rule-based methodology capable of integrating a broad set of parameters to systematically identify influential researchers across the computer science domain.

## 3. Methodology

Guided by insights obtained from the literature review, we designed a comprehensive methodology to derive in- terpretable rules for identifying high-impact researchers within the computer science domain. To pinpoint the most influential quantitative indicators, we utilized a Multi-Layer Perceptron (MLP) model in conjunction with Recursive Feature Elimination (RFE), effectively leveraging a broad spectrum of bibliometric features. The com- plete workflow, illustrated in Figure 1, is structured into five sequential stages: (i) selecting the relevant domain and assembling the dataset, (ii) computing a wide range of author-centric bibliometric metrics, (iii) ranking these metrics based on their predictive relevance, (iv) using the highest-ranked features to classify award recipients, and (v) extracting human-interpretable rules via decision tree algorithms.

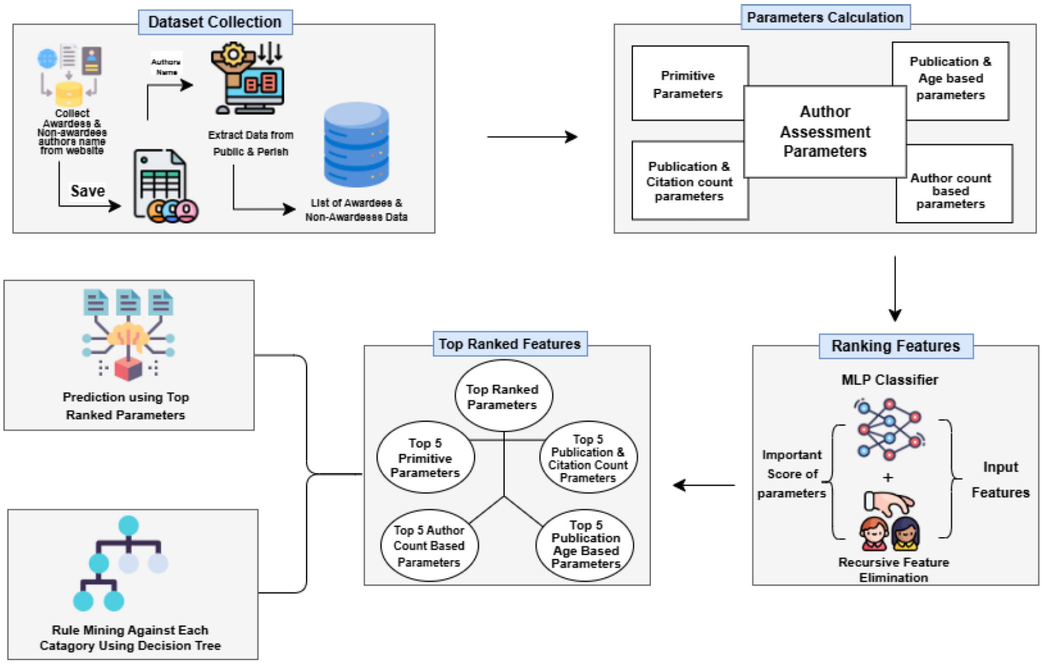


**Figure 1**

**Architecture Diagram**

The subsequent subsections provide a detailed explanation of each individual component.

### 3.1 *Domain Selection and Dataset Collection*

Evaluating the effectiveness of the proposed method required the use of a domain-specific dataset tailored to the context of the study. This study focused on the computer science discipline, chosen for its extensive research legacy, significant impact on technological advancement, and fast-paced development. The field's continual in- novation and relevance in shaping modern solutions make it an ideal candidate for analysis. Additionally, its dynamic characteristics have been highlighted in earlier studies [2, 52].

The dataset consisted of 1,200 entries, equally divided into two groups: 600 awardees and 600 non-awardees. This balanced distribution was chosen to support statistically reliable comparisons and ensure fair representa- tion in experimental analyses. The sample size also corresponds to the accessible and verified awardee profiles collected over the last thirty years.

Information on awardees was obtained from renowned institutions such as the ACM and the IEEE, both of which honor exceptional contributions in research through highly esteemed awards. Data collection covered the years 1990 to 2020 to capture a wide temporal

**Table 1**
**Prestigious Awards Considered for Dataset**

| Award Name | Issuing Organization |
|---|---|
| Turing Award | ACM |
| Computer Pioneer Award | IEEE |
| John von Neumann Medal | IEEE |
| Technical Achievement Award | IEEE |

range and ensure data robustness. The selected awards include prestigious honors such as the ACM Turing Award, Computer Pioneer Award, IEEE John von Neumann Medal, and IEEE Technical Achievement Award, as summarized in Table 1. For the non-awardee group, data was derived from previously utilized datasets, notably those compiled by Samreen et al. [52]. Comprehensive statistics are presented in Table 2.

Information on awardees—including their names and the respective years they received recognition was gath- ered from the official websites of major professional bodies, covering a thirty-year period. To collect detailed researcher profiles, the Publish or Perish application was employed, which utilizes metadata retrieved via Google Scholar's search infrastructure. A "pre-award" data collection approach was adopted to ensure temporal integrity, meaning only publications published before the award year were included in the dataset. To maintain class bal- ance, the number of non-awardees selected per year was matched to the number of awardees. For example, if 15 researchers received awards in the year 1999, then 15 non-awardees were also sampled from prior to that year.

A comprehensive data cleaning process was performed before analysis, particularly focusing on the entries obtained through Google Scholar. This phase was essential for identifying and removing incorrect, duplicate, or unrelated entries that might compromise the validity of results. Data accuracy checks and reduplication procedures were carefully applied. Additionally, two refinement techniques were implemented to improve dataset quality: (i) filtering to retain only publications relevant to the computer science field, and (ii) applying author disambiguation methods to resolve identity confusion an issue frequently encountered in Google Scholar based datasets [53]. The dataset was organized into two distinct groups: awardees and non-awardees. The names of 600 awardees were obtained directly from the official websites of relevant professional societies, thereby removing the need for author disambiguation in this category. In contrast, data for non-awardees were taken from the dataset developed by Samreen et al. [52]. An author disambiguation procedure was then applied, revealing two primary types of ambiguity: (i) instances where both first and last names were identical, requiring close inspection, and (ii) cases involving the same last name but differing first names, which needed additional verification. These were resolved following established techniques found in prior studies [37, 54].

**Table 2**
**Dataset Comprehensive Statistics**

| Researchers Metadata | Count |
|---|---|
| Total No of Authors | 1200 |
| Total No of Awardees | 600 |
| Total No of Non Awardees | 600 |
| Total No of Citation | 32,801,476 |
| Total No of Publication | 171,388 |

Among the 600 non-awardee entries, no instances were found where both the first and last names were exactly the same, rendering the first case irrelevant. However, 33 instances were identified where multiple individuals shared the same last name. Upon further analysis, it was determined that 37 out of 51 such cases referred to different individuals, while 20 out of 49 were confirmed to be name variants of the same author. To preserve balance in the dataset, additional unique non-awardee authors were incorporated accordingly.

The Publish or Perish software played a critical role in resolving these ambiguities. Its robust search features enabled the detection of naming inconsistencies and variations, thereby enhancing both the precision and completeness of the final dataset.

All data collected was publicly available and used strictly for academic research, in accordance with fair-use practices. While the dataset predominantly includes researchers with profiles indexed in Google Scholar, this may introduce some regional or language-related biases that should be considered when interpreting results.

### 3.2 *Assessment Parameter Calculation*

We computed a total of 64 bibliometric indicators for each researcher in the dataset. These were categorized into four groups as follows:

#### 3.2.1 Primitive Parameters

This category encompasses basic bibliometric indicators, including Total Publications, Total Citations, Active Re- search Years, Citations per Year, Citations per Publication, Average Authors per Paper, Citations per Contributor, and Papers per Contributor.

#### 3.2.2 Parameters based on Publication and Citation count

This category encompasses a wide array of bibliometric indices that extend or refine the traditional h-index. These include: the H-index, G-index, E-index, H-core Citations, A-index, R-index, P-index, M-index, F-index, T-index, Q2-index, and Tapered H-index. Additional metrics in this group are Maxprod, Wu-index, Pi-index, Weighted H-index, H(2)-index, Woginger-index, GH-index, RM-index, X-index, HG-index, as well as upper, center, and lower variants of the H2-index. Other notable indicators are the K-dash index, Rational H-index, Real H-index, I10-index, Normalized H-index, K-index, W-index, and the H-dash index.

3.2.3 Parameters based on Author Count

This group comprises authorship-adjusted metrics, including the HI-index, normalized HI, Hm-index, Gm-index, Hf-index, Gf-index, GF-index, K-normalized index, W-normalized index, Pure H-index, as well as fractional variants of the G and H indices, and the normalized HI-index.

3.2.4 Parameters based on Age

These metrics incorporate the time-related aspect of scholarly output and consist of indicators such as the Platinum H-index, M-Quotient, AW-index, AR-index, V-index, Ha-index, Hc-index (also known as the Contemporary H- index), and the Age-Weighted Citation Rate (AWCR).

3.3 *Ranking of Author Assessment Parameters*

Feature ranking is essential in machine learning as it enhances model performance, reduces dimensionality, prevents overfitting, and improves interpretability [55]. Figure 2 depicts the process used to rank author metrics.

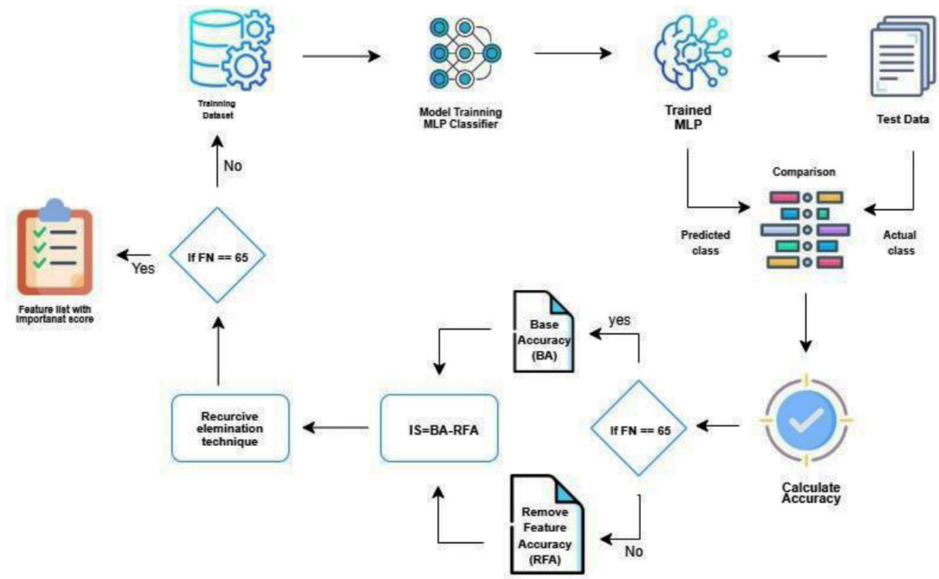


**Figure 2**

**Ranking Process Using MLP and Recursive Elimination**

Our approach involved splitting the dataset into two set with the ratio 80:20 for training and validation. A baseline accuracy was recorded using all 64 features. Each feature was then removed in succession, the model retrained, and the drop in performance used to assign an importance score.

The MLP model employed in this study is a deep feedforward neural network architecture consisting of several hidden layers [56, 57]. During forward propagation, each neuron's output is calculated using:

$$X = WA + b \tag{1}$$

Where $W$ is the weight matrix and $b$ the bias vector. An activation function is applied:

$$C = g(X) \tag{2}$$

We employed ReLU in hidden layers:

$$f(X) = \max(0, X) \tag{3}$$

And Softmax in the output layer:

$$\text{Softmax}(X_i) = \sum_{j=1}^{J} \frac{e^{X_i}}{e^{X_j}} \tag{4}$$

The model's error was calculated using mean squared error:

$$L(z, \hat{z}) = \frac{1}{n} \sum_{i=1}^{n} (z_i - \hat{z}_i)^2 \tag{5}$$

To stabilize training and prevent gradient issues, batch normalization was applied:

$$X_i = \frac{X_i - Mean_i}{StdDev_i} \tag{6}$$

The final classification label was determined using:

$$\text{Output} = \max(\text{Predicted Vector}) \tag{7}$$

We used 10 hidden layers (10 neurons each), ReLU activations, and early stopping (after 40 stagnant epochs).The Adam optimization algorithm was utilized with a learning rate set to 0.0003 and a batch size of 64. Recursive elimination provided refined importance scores to guide rule generation.

### 3.4 *Predictive Modeling with High-Importance Features*

From the ranked results, we selected the top five metrics from each category. These were used to assess researcher ranks and award likelihoods. The outcome was a prioritized list of researchers based on each parameter group's top features.

### 3.5 *Rule Extraction Using the Top 5 Parameters from Each Category*

To derive interpretable rules, we trained decision tree models using the top-ranked metrics per category [58, 59]. Decision trees iteratively split datasets to optimize classification, stopping when a leaf node is pure or further division adds no value.

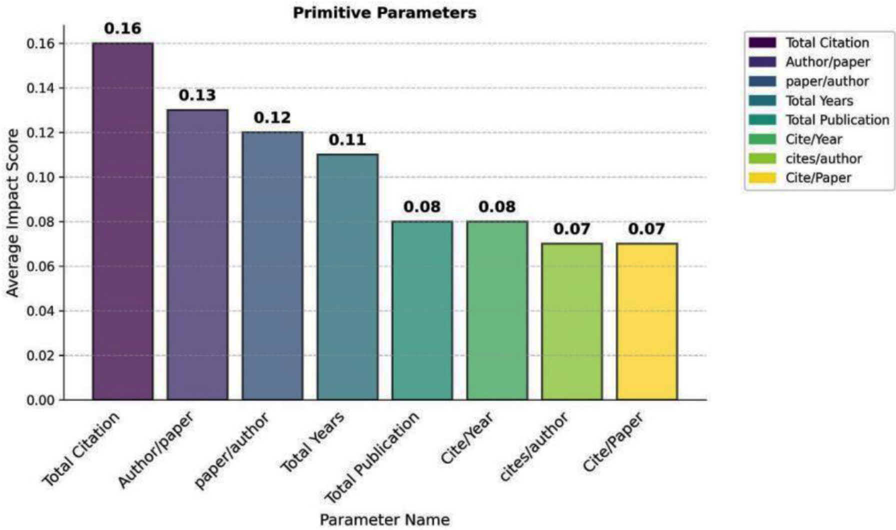


**Figure 3**

**Primitive Parameters**

We used the Gini Index to guide splits, defined as:

$$GI = 1 - \sum_{j=1}^{n} (P_j)^2 \tag{8}$$

Where $P_j$ is the proportion of records in class $j$. A Gini score of 0 implies perfect class purity; 1 indicates total disorder.

Due to the number of features, trees initially grew large and complex. We applied post-pruning to sim- plify them, removing branches with low contribution. A cost-complexity trade-off was used to retain the most generalizable rules, ensuring both interpretability and classification strength.

## 4. Results Discussion

This section presents an in-depth analysis and interpretation of the outcomes derived from applying the proposed methodology to the computer science dataset.

### 4.1 *Ranking of Author Assessment Parameters*

We first present a comprehensive list of the most influential metrics as determined by the combination of Re- cursive Feature Elimination (RFE) and the Multilayer Perceptron (MLP) model. These refined parameters were subsequently used to formulate classification rules via decision tree analysis.

In our analysis of the computer science domain, we assessed the impact and classification effectiveness of parameters organized into distinct categories, as depicted in Figures 3 to 7. Among the primitive metrics, the *Total Citation* indicator showed the greatest average importance with a score of 0.16 and a classification accuracy of 59.83%. The next most impactful parameter was *Author/Paper*, which yielded an average importance of 0.13 and an accuracy of 51.40%. This was followed by *Paper/Author*, which recorded an impact score of 0.12 and an accuracy of 46.35%. *Total Years* came next, with an importance value of 0.11 and 56.74% accuracy. Both *Cites/Year* and *Total Publication* had lower importance scores of 0.08, corresponding to accuracies of 46.35% and 57.87%, respectively.

In the category of publication age-related metrics, the *Platinum h-index* ranked highest with an average im- portance of 0.19 and 64.05% accuracy. Close behind were the *hc index* and *m-quotient index*, each with an impact of 0.17 and accuracies of 61.52% and 62.92%, respectively. The *AW index* also performed competitively, with an average importance of 0.15 and a classification accuracy of 64.33%.

Among the author count-based metrics, the *Normalized hi index* exhibited the strongest influence, achieving an impact score of 0.21 and a classification accuracy of 63.48%. It was followed by the *gF index* with an impact of 0.20 and a classification accuracy of 62.92%. Additional effective metrics included the *gm index* (0.17 impact, 57.87% accuracy), *hi index* (0.12 impact, 62.64%), and *hl norm* (0.12 impact, 63.20%).

Regarding publication and citation-based features, the *Normalized h index* once again proved to be the most significant, showing an average impact of 0.21 and an accuracy of 64.05%. The *h2 upper index* followed with an impact value of 0.18 and 63.76% accuracy. The *h2 center index* contributed a score of 0.16 and 60.67% accuracy. Two additional metrics, *Maxprod* and *Rational h index*, displayed moderate importance (0.14 each) and accuracies of 46.35% and 57.30%, respectively.

When considering all parameter groups together, the top three indicators were the *Normalized h index* (0.21 impact, 64.05% accuracy), *gF index* (0.20, 62.92%), and *Platinum h*

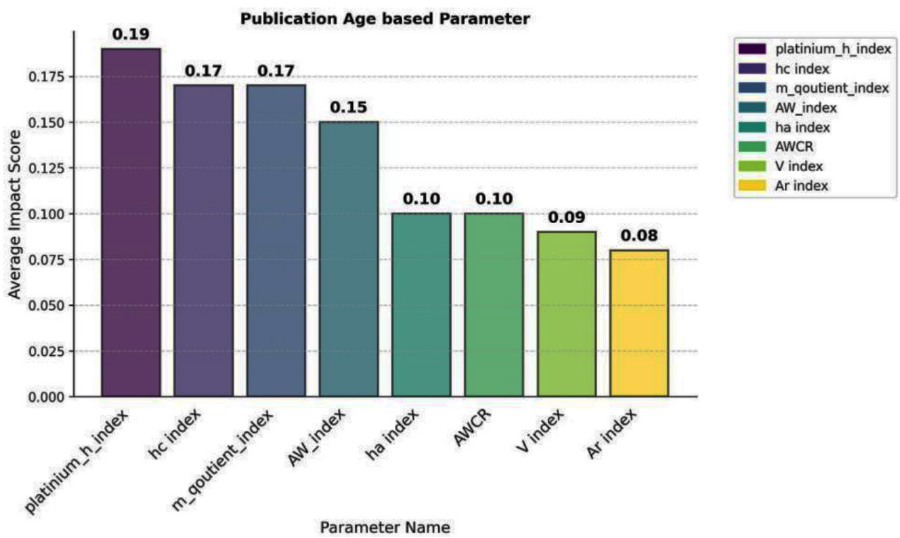


**Figure 4**

**Publication Age-Based Parameters**

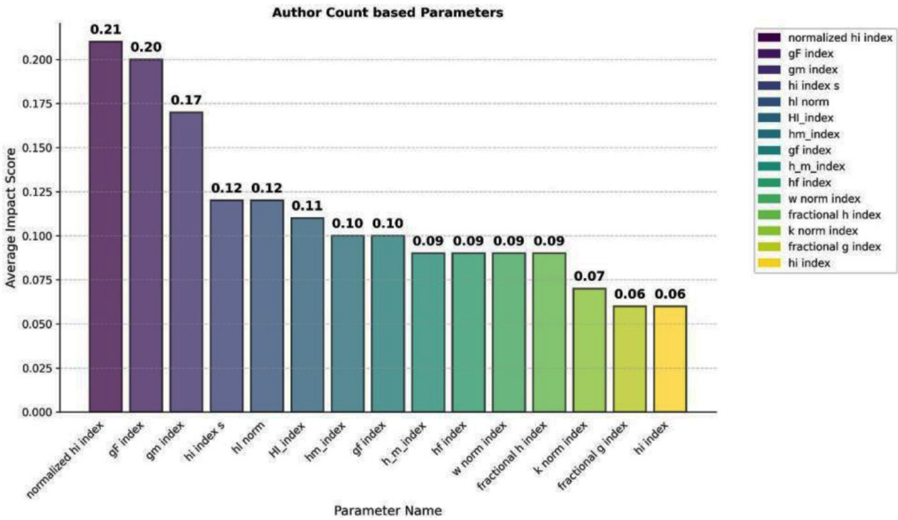


**Figure 5**

**Author Count-Based Parameters**

*index* (0.19, 64.05%). Other high-performing metrics included *h2 upper index*, *gm index*, *hc index*, and *m-quotient index*, all of which exhibited strong impact scores and classification capabilities, reinforcing their usefulness for evaluating scholarly performance in the Computer Science domain.

## 5. Rule Mining based on Decision Tree

This section describes the rule generation procedure conducted for each parameter category, utilizing the top five features identified during the preceding phase. To uncover common patterns among awardees, we utilized decision trees configured with the Gini index as the splitting criterion. For each parameter group, a complete decision tree was generated, resulting in multiple rules. Each rule independently acts as a ranking guideline to help identify award-deserving researchers.

The terminal nodes (leaf nodes) of the tree define the classification result. Rules that lead to a "positive" label (i.e., class 1, denoting award recipients) are visually represented in blue and light blue within the decision tree figures. The rest of the nodes reflect a "negative" classification (class 0). A rule is assigned its class based on the class distribution at its leaf node. For instance, if a rule applies to 50 cases, where 40 belong to class 1 and 10 to class 0, that rule is considered predictive of the positive class. We extracted all such rules and sorted them according to the number of awardees they successfully identified.

It is important to acknowledge a limitation of decision trees: they treat conditions of the form $\leq X$ and $\geq X$ equally, even though this can be illogical when applied to performance metrics. In real-world evaluation,

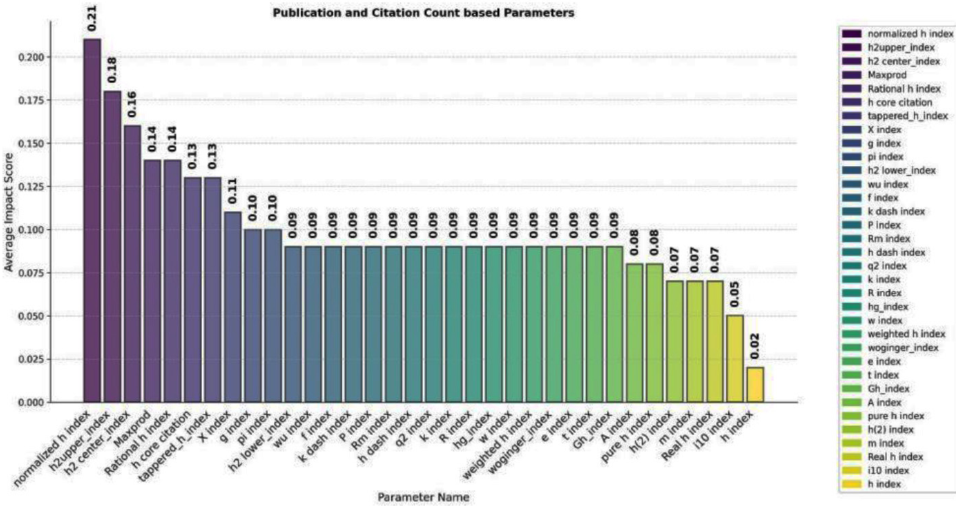


**Figure 6**

**Publication and Citation Count-Based Parameters**

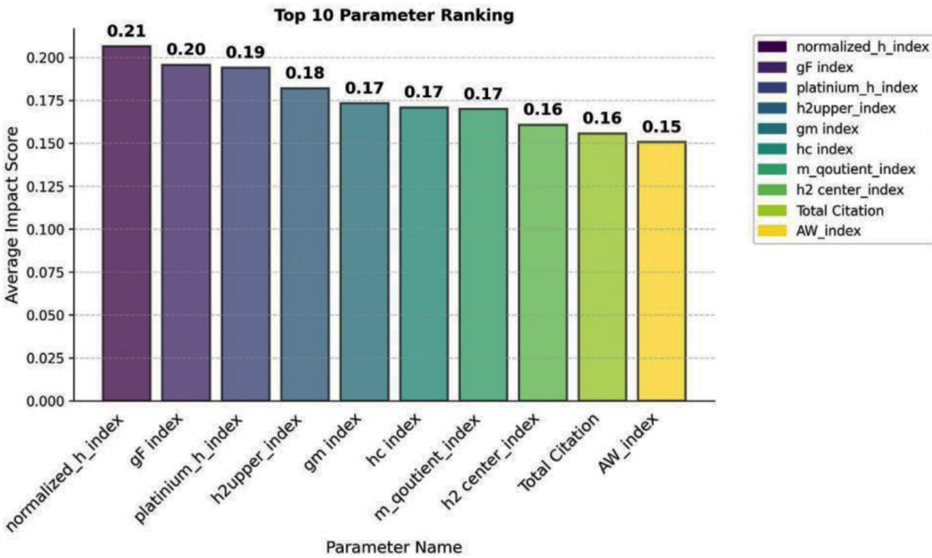


**Figure 7**

**Combined Metrics from All Categories**

imposing an upper limit where exceeding a certain value disqualifies a researcher makes little sense. Since these indicators measure academic achievement, any threshold suggesting a cap on success is inappropriate. Therefore, while interpreting and presenting rules, we disregarded those that imply such counterintuitive boundaries.

To assess the quality of each rule, we employed standard evaluation metrics:

$$\text{Precision} = \frac{\text{Awardees Returned by Rule}}{\text{Total Authors Returned by Rule}} \tag{9}$$

$$\text{Recall} = \frac{\text{Awardees Returned by Rule}}{\text{Total Awardees in Dataset}} \tag{10}$$

$$\text{F - Measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

Based on these metrics, we selected the highest-performing rule for each category to showcase. These top rules are discussed below.

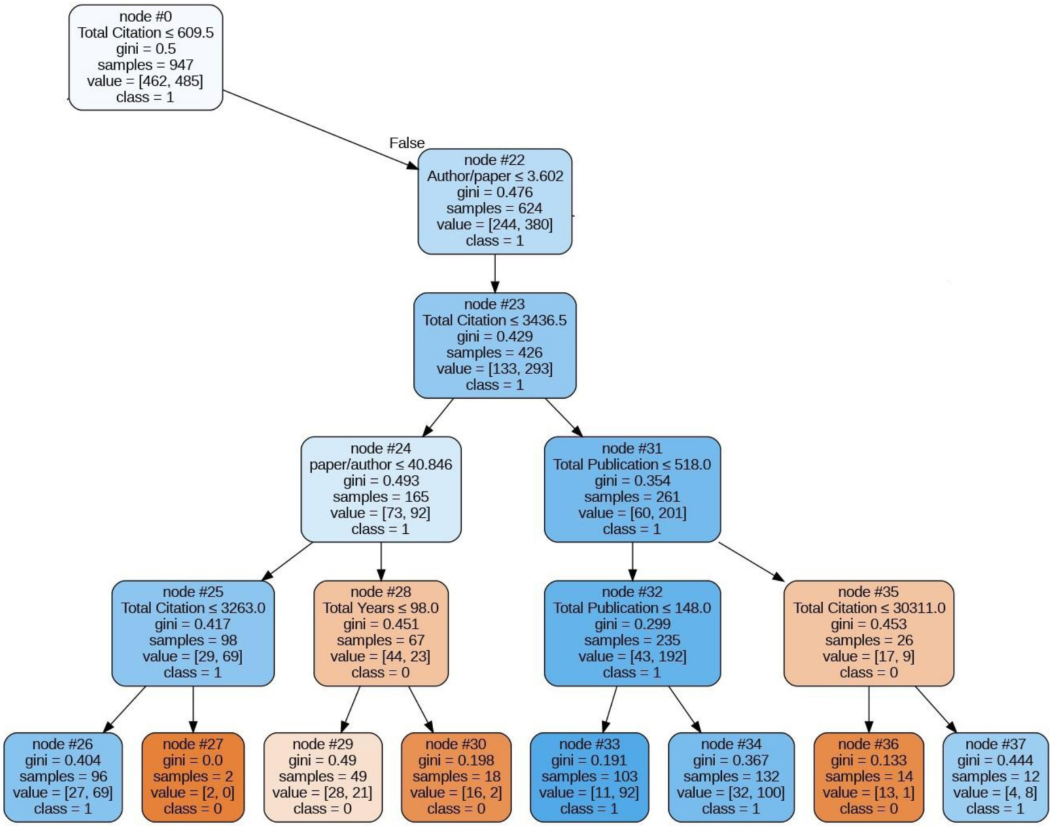


**Figure 8**

**Decision Tree Segment for Primitive Parameter Rule**

5.1 *Rule extracted based on Primitive Parameters*

In this stage of the analysis, we focused on the primitive parameters—basic yet foundational indicators of a re- searcher's scholarly activity. From the comprehensive set of features, five metrics emerged as the highest-ranked based on their contribution to classification accuracy: *Total Publications*, *Total Citations*, *Total Years*, *Paper/Au- thor*, and *Author/Paper*. The first three metrics represent cumulative measures of a researcher's academic output and influence, capturing overall productivity and citation impact accumulated over time. In contrast, *Paper/Author* and *Author/Paper* are fractionalized productivity indicators that reflect co-authorship patterns, offering insight into collaborative intensity and individual contribution per publication.

To extract interpretable patterns from these features, a Decision Tree classifier was trained, aiming to distinguish awardees from non-awardees based solely on these primitive metrics. The classifier achieved a vali- dation accuracy of 64%, indicating that even basic bibliometric information can provide meaningful signals for recognizing scholarly excellence.

The decision tree generated eight rules corresponding to the positive class (i.e., identifying award recipients). Among these, the most effective rule—determined by its highest F-Measure (0.36)—was selected for further analysis. This rule achieved a precision of 0.76 and a recall of 0.21 on the training dataset, suggesting that while it was highly specific in its predictions, it had limited sensitivity. The subtree containing this optimal rule is illustrated in Figure 8.

**Top Rule: IF** Total Citations > 3436.5 **AND** Author/Paper ≤ 3.6 **AND** 148 < Total Publications ≤ 518 ⇒ **Award Recipient**

This rule suggests that a researcher is likely to be recognized with a prestigious award if they meet the fol- lowing conditions: (i) they have accumulated more than 3436.5 total citations, indicating substantial scholarly impact; (ii) the average number of authors per

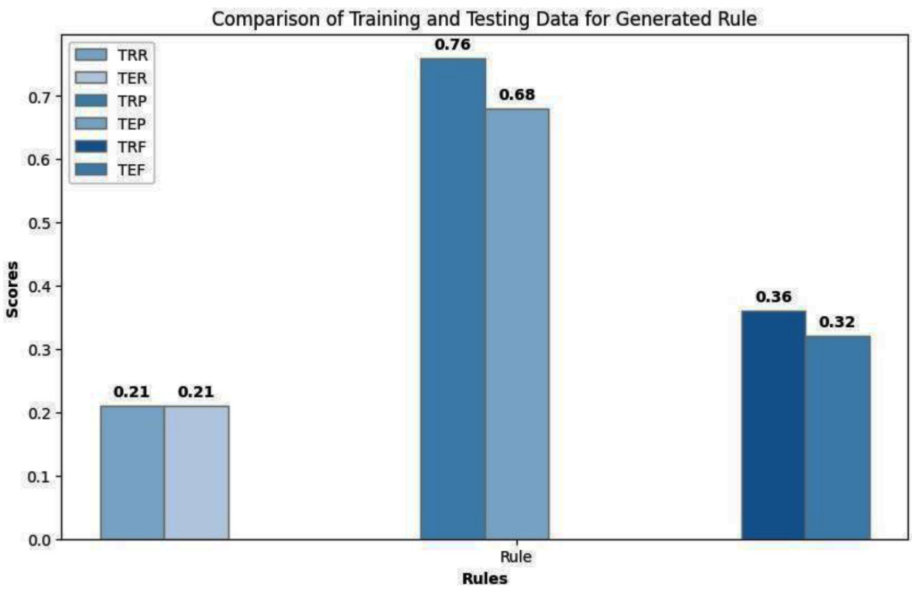


**Figure 9**

**Performance Metrics: Training vs. Testing (Primitive Parameters)**

paper does not exceed 3.6, implying a moderate level of individual contribution; and (iii) their publication count lies between 148 and 518, reflecting a significant, yet realistically achievable, level of productivity. These criteria collectively highlight a profile of sustained, high-quality research output with balanced collaboration.

To assess the generalizability of this rule, it was applied to a separate testing dataset. The performance metrics, summarized in Figure 9, show that the model retained a recall of 0.21, while the precision slightly dropped to

0.68. The resulting F-Measure in the test phase was 0.32, indicating a minor reduction in performance compared to training but affirming the stability and interpretability of the rule across datasets.

The relatively high precision of this rule makes it suitable for high-confidence candidate identification, even though its recall is modest. In contexts where false positives must be minimized—such as nominating individuals for awards or grants—this rule provides a valuable, interpretable heuristic for decision-makers.

5.2 *Rule extracted based on Author Count-Based Parameters*

This section examines the predictive power of metrics that adjust for co-authorship, offering a more refined view of individual research contributions. The top five features selected from this category include: *Normalized h-index*, *gF index*, *hi index*, *gm index*, and *hi norm*. These indicators are designed to account for the dilution of credit in multi-author papers, thereby capturing a more accurate representation of a researcher's independent scholarly influence.

A decision tree trained using these five metrics produced a classification accuracy of 62%, indicating a moderate yet meaningful ability to discriminate between awardees and non-awardees based solely on author count-adjusted indicators. From the resulting tree structure, nine distinct rules associated with the positive class (award recipients) were identified. The most effective among these achieved an F-Measure of 0.48, with a recall of 0.35 and precision of 0.78, suggesting that this rule offers both reasonable sensitivity and high confidence in prediction. The decision tree segment highlighting this rule is illustrated in Figure 10.

**Top Rule: IF** Normalized h-index > 0.2 **AND** 5.5 < hi norm ≤ 91.5 **AND** hi index > 5.9 **AND** gm index ≤ 98.5 ⇒ **Award Recipient**

This rule indicates that researchers who maintain a high *Normalized h-index*—signifying consistent impact despite shared authorship—along with strong individual performance indicators such as *hi index* and *hi norm*, are more likely to be recognized. Meanwhile, a constrained *gm index* reflects a balanced authorship pattern, avoiding extremes that could inflate perceived contribution.

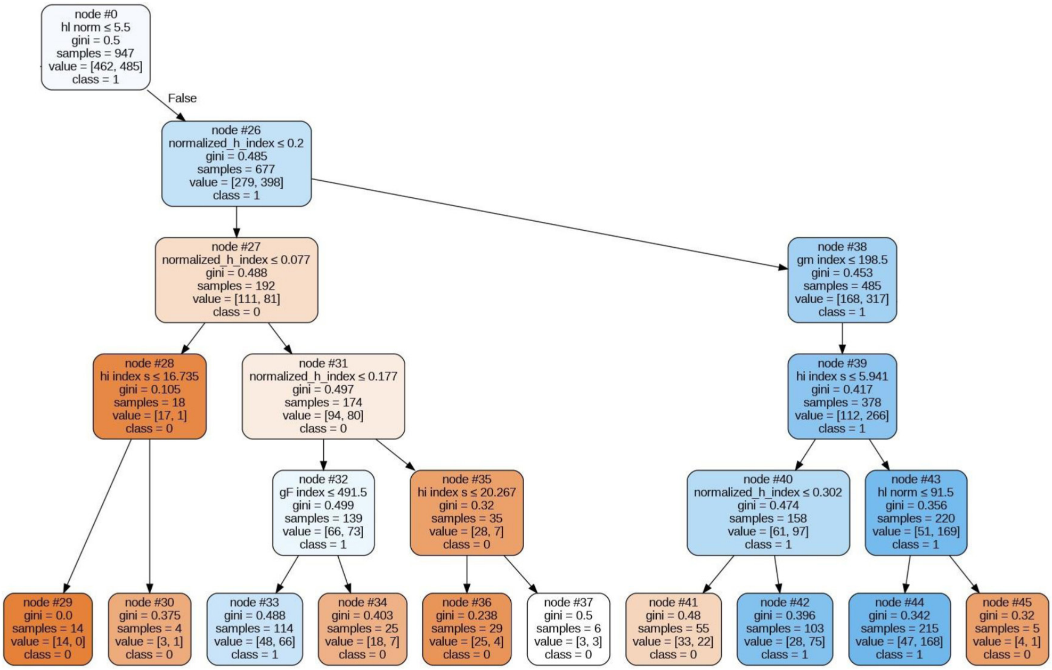


**Figure 10**

**Decision Tree Segment for Author Count-Based Rule**

To evaluate its generalizability, this rule was applied to an independent test set. The performance remained relatively strong, with a recall of 0.29, precision of 0.72, and an F-Measure of 0.42. Although slightly lower than the training performance, the rule demonstrated consistent predictive utility across datasets, as shown in Figure 11. These findings underscore the value of author count-sensitive metrics in understanding individual researcher impact. By filtering out inflation effects caused by high co-authorship, the derived rules offer a more equitable and interpretable basis for identifying deserving candidates for academic recognition.

5.3 *Rule extracted based on Age-Based Parameters*

Age-sensitive metrics aim to account for the influence of a researcher's career duration and publication history on scholarly impact. These indicators adjust for temporal factors, enabling fairer comparisons between early-career and senior researchers. The top five metrics identified in this category were: *Platinum h-index*, *AW index*, *M- Quotient*, *hc index*, and *ha index*. Each of these parameters provides a nuanced lens into the dynamics of academic productivity over time.

A decision tree was trained using these features, yielding a classification accuracy of 62%. From the generated tree, nine rules relevant to the positive class (i.e., award recipients) were extracted. Among them, the most promi- nent rule demonstrated an F-Measure of

0.60, with a recall of 0.54 and a precision of 0.67. This rule was selected for further interpretation due to its strong balance between sensitivity and specificity. The relevant segment of the decision tree is presented in Figure 12.

**Top Rule: IF** 5.8 < Platinum h-index ≤ 2024.3 **AND** AW index > 7.6 **AND** M-Quotient ≤ 0.445 ⇒ **Award Recipient**

This rule reveals that awardees often exhibit high long-term impact (reflected by a substantial *Platinum h- index*), consistent productivity over time (indicated by a strong *AW index*), and a moderate *M-Quotient*, which tempers productivity by career length. These values collectively suggest that impactful researchers maintain steady, quality output across their academic careers without depending on a short burst of publications.

To test the robustness of this rule, it was validated on a hold-out dataset. As shown in Figure 13, the model pre- served a recall of 0.54, while the precision experienced a minor decrease to 0.62. This resulted in a slightly lower F-Measure of 0.58. Despite the dip, the rule retained strong predictive relevance, confirming its generalizability across unseen data.

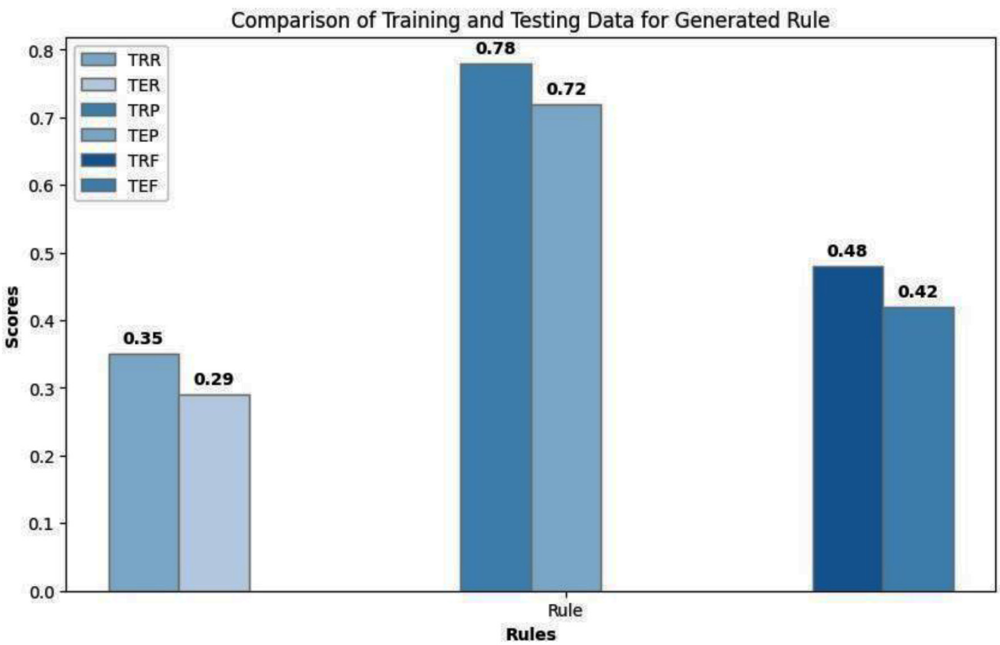


**Figure 11**

**Performance Metrics: Training vs. Testing (Author Count-Based)**

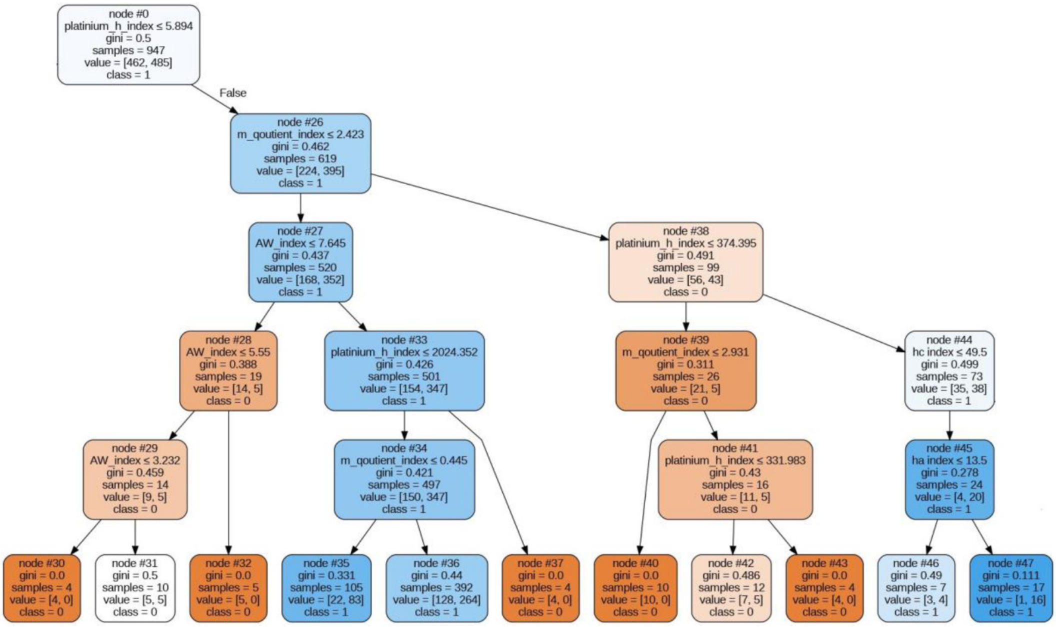


**Figure 12**

**Decision Tree Segment for Age-Based Rule**

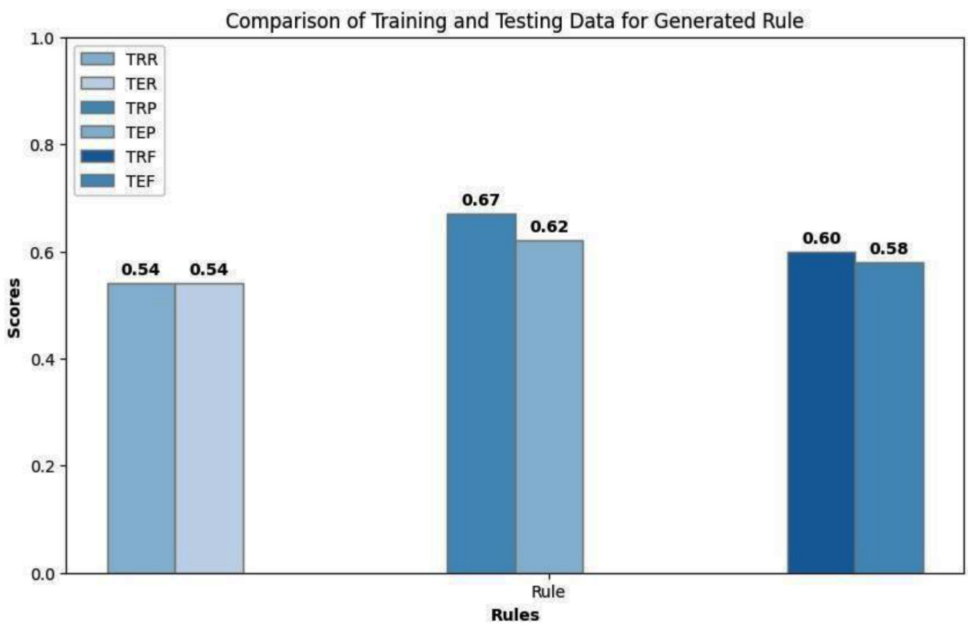


**Figure 13**

**Performance Metrics: Training vs. Testing (Age-Based)**

These findings underscore the importance of integrating time-adjusted metrics when assessing research per- formance. Age-sensitive indicators mitigate biases associated with career length, offering a more equitable basis for recognizing scholarly excellence across different stages of academic progression.

5.4 *Rule extracted based on the Publication and Citation Count-Based Parameters*

This category comprises metrics designed to capture the magnitude and structure of a researcher's citation record. These parameters not only quantify raw impact but also assess how citations are distributed across a scholar's body of work. The five most influential features identified within this group were: *Maxprod*, *h2 upper index*, *h2 center index*, *normalized h-index*, and *rational h-index*. These indicators provide deeper insight into citation dy- namics, particularly emphasizing productivity depth, the concentration of influential publications, and normalized performance.

The decision tree constructed using these variables achieved a classification accuracy of 65%, which was the highest among all feature groups examined in this study. From the generated model, eleven distinct decision rules were associated with the awardee class. Among these, the rule that yielded the best performance—based on the F-Measure—was selected for interpretation. This rule achieved a recall of 0.42 and a notably high precision of 0.87, resulting in an F-Measure of 0.57 during training. The relevant subtree is visualized in Figure 14.

**Top Rule: IF** Maxprod > 482.0 **AND** h2 upper index > 59.5 **AND** normalized h-index ≤ 0.9 **AND** h2 center index ≤ 22.9 ⇒ **Award Recipient**

This rule identifies a researcher as a potential award recipient if they meet the following conditions: (i) a *Maxprod* score exceeding 482.0, indicating a high-productivity paper with substantial citations; (ii) an *h2 upper index* above 59.5, highlighting strong citation depth in top-tier publications; (iii) a moderate *normalized h-index* (no greater than 0.9), suggesting

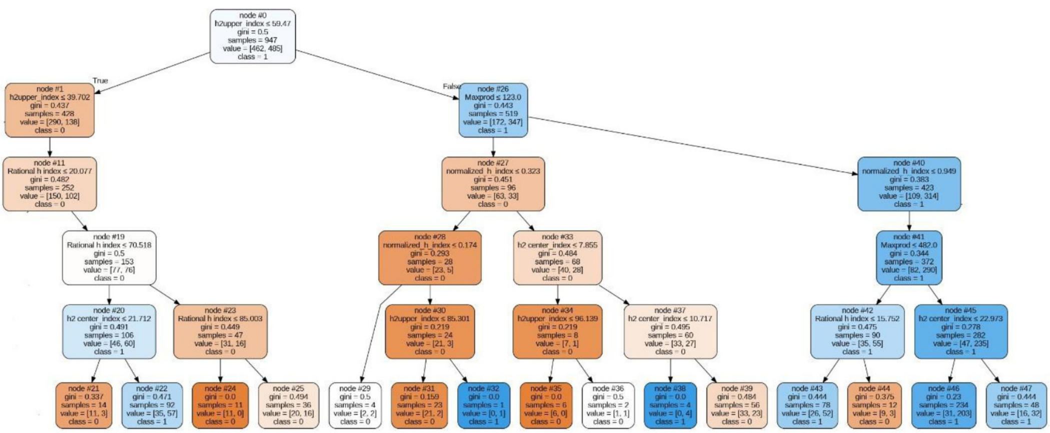


**Figure 14**

**Decision Tree Segment for Publication and Citation Count-Based Rule**

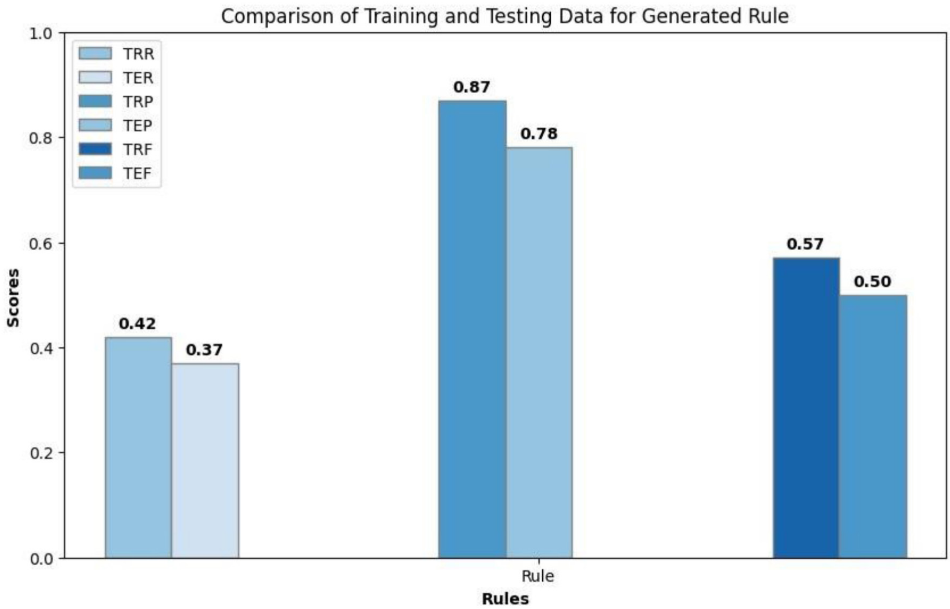


**Figure 15**

**Performance Metrics: Training vs. Testing (Publication and Citation Count-Based)**

balanced impact accounting for publication volume; and (iv) an *h2 center index* not exceeding 22.9, reflecting a well-distributed core impact. Notably, this rule demonstrates that exceptionally high values in a few critical metrics can outweigh relatively moderate values in others when distinguishing highly impactful scholars.

To evaluate the reliability of this rule, it was tested on a separate validation set. As shown in Figure 15, the recall slightly declined to 0.37 and precision to 0.78, resulting in an F-Measure of 0.50. While this indicates a modest decrease in predictive performance compared to training, the results confirm the rule's generalizability and robustness across unseen data.

The rule underscores the value of advanced citation-based indicators in identifying top researchers. It demon- strates that sophisticated metrics such as *Maxprod* and layered h2-index variants can capture key patterns of influence often missed by traditional indicators, making them valuable tools in award prediction and academic profiling.

### 5.5 *Combine All Parameters*

In the final phase of our analysis, we integrated the top-performing metrics from each of the four parameter categories—primitive, publication/citation-based, author-count-based, and age-based. This holistic configuration aimed to capture a multi-dimensional view of researcher performance by blending diverse bibliometric dimen- sions. From this combined feature set, five metrics emerged as the most influential: *h2 upper index*, *normalized*

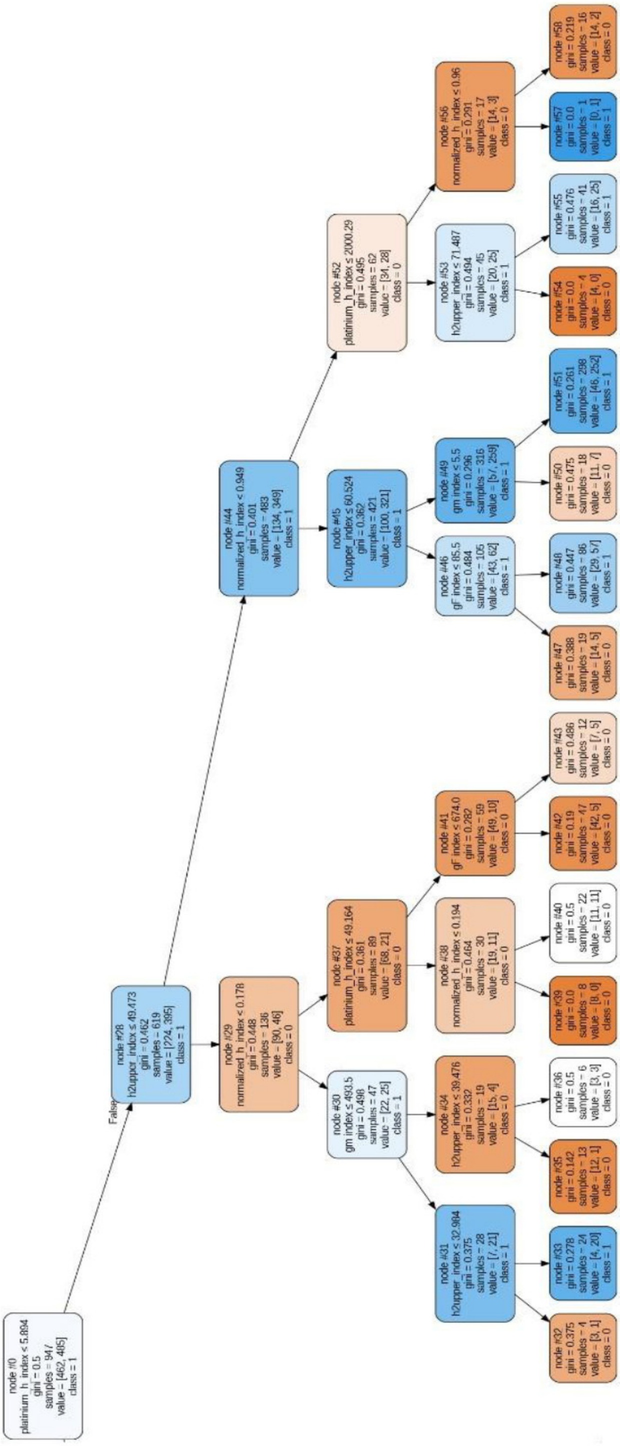


**Figure 16**
**Decision Tree Segment for Combined Parameters Rule**

*h-index*, *platinum h-index*, *gF index*, and *gm index*. These indicators collectively represent a balance of citation depth, co-authorship sensitivity, and temporal adjustment.

A decision tree model trained on this enriched feature space achieved an overall classification accuracy of 67%, outperforming the models based on individual categories. From the resulting decision structure, ten rules that led to the awardee class were identified. The most effective rule achieved an F-Measure of 0.64 on the training set, with a precision of 0.85 and a recall of 0.52. This rule demonstrates both strong confidence and moderate sensitivity in identifying high-impact researchers. The relevant tree segment is shown in Figure 16.

**Top Rule: IF** Platinum h-index > 5.9 **AND** h2 upper index > 60.5 **AND** normalized h-index ≤ 0.9 **AND** gm index > 5.5 ⇒ **Award Recipient**

This rule highlights that individuals with consistently high values in long-term citation indices—such as the *platinum h-index* and *h2 upper index*—combined with a strong *gm index* (accounting for authorship patterns), are highly likely to be recognized through formal awards. Interestingly, a normalized h-index value at or below 0.9, often viewed as moderate, does not preclude recognition if other impactful metrics compensate for it. This reinforces the value of combining multiple metrics to uncover award-worthy profiles that might be overlooked when relying on a single indicator.

To evaluate the rule's effectiveness on unseen data, it was validated on a separate test set. As illustrated in Figure 17, the model experienced a slight decline in recall (to 0.46) and precision (to 0.80), resulting in an F- Measure of 0.59. Despite this decrease, the rule maintained solid performance, demonstrating its generalizability and practical utility across different data partitions.

Overall, the integration of top indicators across multiple bibliometric dimensions enhanced the model's robust- ness and interpretability. This approach not only improved prediction accuracy but also allowed for the derivation of comprehensive rules that reflect both productivity and scholarly influence in a nuanced and balanced manner.

## 6. Discussion on Generated Rules

Table 3 compiles the best rules derived for each parameter group. In the construction of these rules, we excluded any conditions that involved upper limits (e.g., "parameter ≤ X") unless they were practically reasonable. This ensures that all recommended rules reflect logical thresholds that promote achievement, not restrict it. The the- oretical novelty of our approach lies in the integration of 64 distinct parameters from four categories, utilizing

machine learning techniques such as Recursive Feature Elimination (RFE) with a Multilayer Perceptron (MLP) classifier. These data-driven techniques enabled the creation of interpretable, category-specific rules that offer actionable insights for researchers. Unlike traditional indices like the h-index, our method uses a combination of multiple, ranked parameters to derive thresholds that are both logical and motivational, providing a clearer framework for research evaluation.


**Figure 17**

**Performance Metrics: Training vs. Testing (Combined Parameters)**

Additionally, our method's practical uniqueness lies in the clarity and applicability of these rules in real-world scenarios. The rules derived from our decision tree model offer clear benchmarks for researchers to track their progress toward recognition, providing actionable guidelines for those aspiring to be included in prestigious award lists. The approach shifts from relying on a single indicator to utilizing a composite set of parameters, making it more comprehensive and grounded in empirical evidence from the dataset. This multi-parameter framework can be adapted and applied across various research domains to evaluate impactful researchers and identify those deserving recognition.

The thresholds presented in the rules, such as 46% of awardees exceeding the Cite/Author threshold and 43% meeting the HI-index, Normalized h-index, and gF index thresholds, represent the proportion of awardees who meet at least one threshold from each parameter category. These results illustrate the patterns of characteristics shared by awardees and are not intended to imply that researchers need to meet all thresholds of a single rule to be recognized. Instead, the thresholds serve as guidelines, and a researcher who meets a combination of parameters (e.g., high Cite/Author and HI-index score) would be classified as an awardee. This highlights the flexibility of our framework, where multiple combinations of parameters can lead to award recognition, making it adaptable to different academic fields and disciplines.

Moreover, the percentages reflect the interactions among the parameters, indicating that while not all awardees meet the specific threshold of each parameter, they often exceed the thresholds of other important parameters. Thus, the findings highlight that meeting thresholds from multiple categories is more influential in determining awardees, rather than the fulfillment of a single rule.

Based on the insights gained from Table 3, we propose the following recommendations for researchers seeking to enhance their prospects of formal recognition:

- Aim for a Cite/Author ratio exceeding 1561.3 and a Paper/Author ratio above 10.1, as nearly half of all awardees meet these benchmarks.
- Maintain an HI-index above 6.913, a Normalized h-index greater than 0.461, and a gF index exceeding 24.5.

These metrics correspond to more than 40% of the awardee group.

- Work toward achieving a Platinum h-index above 37.6, a mark achieved by over 21% of awardees.

**Table 3**

**Rules Based on Top Features from Each Parameter Category**

| No. | Rule Description | Category | P | R | F-M |
|---|---|---|---|---|---|
| 1 | IF Cite/Author > 1561.3 AND Paper/Author > 10.1 → Award Recipient | Primitive | 0.67 | 0.34 | 0.46 |
| 2 | IF Normalized h-index > 0.461 AND HI-index > 3.88 AND gF index > 24.5 → Award Recipient | Author Count | 0.28 | 0.97 | 0.43 |
| 3 | IF Platinum h-index > 37.6 → Award Recipient | Age-Based | 0.12 | 0.72 | 0.21 |
| 4 | IF Maxprod ≥ 188.5 → Award Recipient | Pub/Cit Count | 0.85 | 0.63 | 0.73 |
| 5 | IF Normalized h-index > 0.461 AND HI-index ≥ 6.913 → Award Recipient | Combined | 0.29 | 1.00 | 0.45 |

- Consider optimizing for a Maxprod score of at least 188.5, which aligns with the profiles of approximately 73% of recipients.

It is also worth noting that several parameters were deliberately excluded from the final rule set—such as Total Publications, Cite/Paper, AW index, X index, and f index—because the decision tree logic required these values to fall below a certain threshold, a condition not considered relevant or desirable in practical evaluation. The exclusion of such thresholds underscores our commitment to creating rules that enable achievement, rather than constrain progress. This emphasis on achievement-based criteria is central to the practical applicability and novelty of our framework.

While our framework identifies the most significant parameters for award recognition, it is important to acknowledge that other qualitative factors—such as collaboration networks, funding sources, and domain-specific characteristics—could also play a role in a researcher's recognition but were not modeled in this study. Our rules are based on quantitative data and provide actionable guidelines for researchers. However, we recognize that addi- tional, unquantified variables may influence the final recognition outcome. This aspect could be explored in future work, where qualitative and quantitative factors can be integrated into a more holistic framework for evaluating researcher impact.

By combining established quantitative parameters in a novel way, our framework offers a unique, data-driven approach to researcher evaluation. This framework provides practical value not only to individual researchers striving for recognition but also to academic institutions and funding bodies seeking more objective and trans- parent recognition practices. Moreover, given its ability to handle multiple parameters, our approach has a high degree of transferability across academic fields and can be adapted for use in domains outside of computer sci- ence, making it a generalizable tool for impactful researcher evaluation. Future work should test the framework's adaptability and effectiveness in other domains to confirm its transferability and further enhance its practical application.

## 7. Concluding Remarks

In this research we introduce a novel framework for evaluating impactful researchers in the computer science do- main, leveraging 64 distinct bibliometric parameters from four categories: primitive metrics, publication/citation- based parameters, author-count-based metrics, and age-weighted indicators. The findings highlight the significant role of traditional measures like total citation count and normalized h-index, while also emphasizing the effective- ness of new parameters, such as the Platinum h-index and Normalized h-index, in distinguishing award-worthy researchers.

By employing machine learning techniques like Recursive Feature Elimination (RFE) and Multilayer Per- ceptron (MLP), we identified the most influential parameters and developed a set of interpretable decision rules. These rules offer a robust, multi-dimensional approach to researcher evaluation, moving beyond the limitations of single-metric assessments. The framework provides actionable guidelines for both academic institutions and individual researchers, facilitating more objective recognition of scholarly contributions.

Looking ahead, we plan to extend this methodology across multiple academic disciplines to assess its gener- alizability. Future work will also focus on the development of novel indices that incorporate qualitative metrics such as collaboration networks and sustained citation growth which are often overlooked by traditional mea- sures. Additionally, we aim to explore the use of advanced deep learning models to enhance the precision and computational efficiency of researcher recognition systems.

As part of our long-term goals, we intend to integrate emerging metrics like the Q-factor to create a more comprehensive, context-aware evaluation framework. This will involve expanding our dataset to include a broader range of variables, ensuring a more holistic assessment of a researcher's academic contributions.

In conclusion, this study provides a solid foundation for refining the process of academic recognition, offer- ing a flexible and scalable model that can be adapted to various research domains, with the potential for future improvements in both methodology and application. Additionally, there are emerging indicators worth integrating into our evaluation framework. One such metric is the Q-factor proposed by Sinatra et al. [60], which cap- tures variables like mentoring, project involvement, funding success, and overall research trajectory—dimensions often excluded from conventional indices. Incorporating such qualitative factors will require assembling a more comprehensive and multidisciplinary dataset.

Looking ahead, we also plan to explore critical questions, including: Can integrating subjective and objective indicators enhance the predictive capabilities of researcher evaluation models? To what extent can these parame- ters be practically implemented in digital repositories and academic search engines? And finally, is there a need for new metrics that better support nuanced and context-aware decision-making in Scientometrics?

### 7.1 *Data Availability*

The data used in this study is publicly accessible and can be found at the following GitHub repository: https://github.com/ghulammustafacomsat/computerscience.git. This repository contains the compiled dataset of awardees and non-awardees, as well as the computed author assessment parameters utilized in the analysis. The authors affirm that all relevant data supporting the findings are available in this repository.

**Use of AI Tools:** During manuscript preparation, OpenAI's ChatGPT was employed to support language refinement and improve readability. Grammarly was additionally used for grammar and style checks.

## References


[1] G. Mustafa, M. Usman, L. Yu, M. T. Afzal, M. Sulaiman, and A. Shahid, "Multi-label classification of research articles using word2vec and identification of similarity threshold," *Scientific Reports*, vol. 11, no. 1, p. 21900 (2021), doi: 10.1038/s41598-021-01460-7.

[2] G. Mustafa, A. Rauf, A. S. Al-Shamayleh, M. Sulaiman, W. Alrawagfeh, M. T. Afzal, and A. Akhunzada, "Optimizing document classification: Unleashing the power of genetic algorithms," *IEEE Access* (2023), doi: 10.1109/ACCESS.2023.3292248.

[3] A. Sidiropoulos, D. Katsaros, and Y. Manolopoulos, "Generalized Hirsch h-index for disclosing latent facts in citation networks," *Scientometrics*, vol. 72, pp. 253–280 (2007), doi: 10.1007/s11192-007-1722-z.

[4] S. Yang, Q. Yuan, J. Dong, "Are scientometrics, informetrics, and bibliometrics different?" *Data Science and Informetrics*, vol. 1, no. 1, p. 50 (2020).

[5] C. Leibel and L. Bornmann, "What do we know about the disruption index in scientometrics? An overview of the literature," *Scientometrics*, vol. 129, no. 1, pp. 601–639 (2024).

[6] M. Usman, G. Mustafa, and M. T. Afzal, "Ranking of author assessment parameters using logistic regression," *Scientometrics*, vol. 126, no. 1, pp. 335–353 (2021), doi: 10.1007/s11192-020-03769-y.

[7] A. A. Alshdadi, M. Usman, M. O. Alassafi, M. T. Afzal, and R. AlGhamdi, "Formulation of rules for the scientific community using deep learning," *Scientometrics*, vol. 128, no. 3, pp. 1825–1852 (2023), doi: 10.1007/s11192-023-04633-5.

[8] S. M. A. H. Shah, A. Ullah, J. Iqbal, S. Bourouis, S. S. Ullah, S. Hussain, M. Q. Khan, Y. A. Shah, and G. Mustafa, "Classifying and localizing abnormalities in brain MRI using channel attention based semi-Bayesian ensemble voting mechanism and convolutional autoencoder," *IEEE Access*, vol. 11, pp. 75528–75545 (2023).

[9] J. E. Hirsch, "An index to quantify an individual's scientific research output," *Proceedings of the National Academy of Sciences*, vol. 102, no. 46, pp. 16569–16572 (2005), doi: 10.1073/pnas.050765510.

[10] W. Hood and C. Wilson, "The literature of bibliometrics, scientometrics, and informetrics," *Scientometrics*, vol. 52, no. 2, pp. 291–314 (2004).

[11] K. R. Dienes, "Completing h," *Journal of Informetrics*, vol. 9, no. 2, pp. 385–397 (2015). https://doi.org/10.1016/j.joi.2015.01.003

[12] G. Mustafa, A. Rauf, B. Ahmed, M. T. Afzal, A. Akhunzada, and S. Z. Alharthi, "Comprehensive evaluation of publication and citation metrics for quantifying scholarly influence," *IEEE Access*, vol. 11, pp. 65759–65774 (2023).

[13] A. Bihari, S. Tripathi, and A. Deepak, "A review on h-index and its alternative indices," *Journal of Information Science*, vol. 49, no. 3, pp. 624–665 (2023). https://doi.org/10.1177/01655515211014478

[14] G. Mustafa, M. Usman, M. T. Afzal, A. Shahid, and A. Koubaa, "A comprehensive evaluation of metadata-based features to classify research paper's topics," *IEEE Access*, vol. 9, pp. 133500–133509 (2021).

[15] S. Ayaz and M. T. Afzal, "Identification of conversion factor for completing-h index for the field of mathematics," *Scientometrics*, vol. 109, no. 3, pp. 1511–1524 (2016). https://doi.org/10.1007/s11192-016-2122-z

[16] Q. L. Burrell, "On the h-index, the size of the Hirsch core and Jin's a-index," *Journal of Informetrics*, vol. 1, no. 2, pp. 170–177 (2007). https://doi.org/10.1016/j.joi.2007.01.003

[17] B. Jin, L. Liang, R. Rousseau, and L. Egghe, "The r- and ar-indices: Complementing the h-index," *Chinese Science Bulletin*, vol. 52, no. 6, pp. 855–863 (2007). https://doi.org/10.1007/s11434-007-0145-9

[18] N. R. Khan, C. J. Thompson, D. R. Taylor, K. S. Gabrick, A. F. Choudhri, F. R. Boop, and P. Klimo Jr., "Part II: Should the h-index be modified? An analysis of the m-quotient, contemporary h-index, authorship value, and impact factor," *World Neurosurgery*, vol. 80, no. 6, pp. 766–774 (2013). https://doi.org/10.1016/j.wneu.2013.07.011

[19] K. Garg, R. K. Singh, "A comparative study of the bibliometric characteristics of COLLNET Journal of Scientometrics and Information Management and Journal of Scientometric Research," *Journal of Scientometric Research*, vol. 13, no. 2, pp. 636–650 (2024).

[20] Ain, Qurat-ul, Hira Riaz, and Muhammad Tanvir Afzal. "Evaluation of h-index and its citation intensity based variants in the field of mathematics." Scientometrics 119, no. 1: 187-211 (2019).

[21] G. Mustafa, A. Rauf, A. S. Al-Shamayleh, B. Ahmed, W. Alrawagfeh, M. T. Afzal, and A. Akhunzada, "Exploring the significance of publication-age-based parameters for evaluating researcher impact," *IEEE Access*, vol. 11, pp. 86597–86610 (2023).

[22] B. Ahmed, L. Wang, A. S. Al-Shamayleh, M. T. Afzal, G. Mustafa, W. Alrawagfeh, and A. Akhunzada, "Machine learning approach for effective ranking of researcher assessment parameters," *IEEE Access*, vol. 11, pp. 133294–133312 (2023).

[23] N. De, S. M. A. Nayeem, and A. Pal, "F-index of some graph operations," *Discrete Mathematics, Algorithms and Applications*, vol. 8, no. 02, 1650025 (2016). https://doi.org/10.1142/S1793830916500257

[24] M. Schreiber, "A modification of the h-index: The hm-index accounts for multi-authored manuscripts," *Journal of Informetrics*, vol. 2, no. 3, pp. 211–216 (2008). https://doi.org/10.1016/j.joi.2008.05.001

[25] R. Ghani, F. Qayyum, M. T. Afzal, and H. Maurer, "Comprehensive evaluation of h-index and its extensions in the domain of mathematics," *Scientometrics*, vol. 118, pp. 809–822 (2019). https://doi.org/10.1007/s11192-019-03007-0

[26] C. Moreira, P. Calado, and B. Martins, "Learning to rank academic experts in the DBLP dataset," *Expert Systems*, vol. 32, no. 4, pp. 477–493 (2015). https://doi.org/10.1111/exsy.12062

[27] H. Kumar, "About scientometrics: An overview," *Journal of Library and Information Science*, vol. 7, no. 1 (2025).

[28] S. Jose and J. Franklin, "A study of Indian theses on applied bibliometrics, informetrics, scientometrics and citation analysis," *Annals of Library and Information Studies*, vol. 72, no. 2, pp. 186–200 (2025).

[29] B. Ahmed, W. Li, G. Mustafa, M. T. Afzal, S. Z. Alharthi, and A. Akhunzada, "Evaluating the effectiveness of author-count based metrics in measuring scientific contributions," *IEEE Access* (2023). https://doi.org/10.1109/ACCESS.2023.3309416

[30] R. Crowder, G. Hughes, and W. Hall, "An agent-based approach to finding expertise," in *International Conference on Practical Aspects of Knowledge Management*, pp. 179–188 (2002). https://doi.org/10.1007/3-540-36277-0_17

[31] M. Raheel, S. Ayaz, and M. T. Afzal, "Evaluation of h-index, its variants and extensions based on publication age & citation intensity in civil engineering," *Scientometrics*, vol. 114, pp. 1107–1127 (2018). https://doi.org/10.1007/s11192-017-2633-2

[32] G. Mustafa, A. Rauf, A. S. Al-Shamayleh, M. T. Afzal, A. Waqas, and A. Akhunzada, "Defining quantitative rules for identifying influential researchers: Insights from mathematics domain," *Heliyon*, vol. 10, no. 9 (2024).

[33] G. Mustafa, A. Rauf, and M. T. Afzal, "GK index: Bridging GF and K indices for comprehensive author evaluation," *Knowledge and Information Systems*, vol. 66, no. 9, pp. 5203–5238 (2024).

[34] Q. Burrell, "Hirsch index or Hirsch rate? Some thoughts arising from Liang's data," *Scientometrics*, vol. 73, no. 1, pp. 19–28 (2007). https://doi.org/10.1007/s11192-006-1774-5

[35] R. Liang and X. Jiang, "Scientific ranking over heterogeneous academic hypernetwork," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30 (2016). https://doi.org/10.1609/aaai.v30i1.10004

[36] J. Bar-Ilan, "Informetrics at the beginning of the 21st century—A review," *Journal of Informetrics*, vol. 2, no. 1, pp. 1–52 (2008).

[37] J. Joseph, J. Jose, A. S. Jose, G. G. Ettaniyil, and S. V. Nair, "A scientometric analysis of bibliotherapy: Mapping the research landscape," *Library Hi Tech*, vol. 43, no. 2/3, pp. 1014–1034 (2025).

[38] G. Mustafa, A. Rauf, and M. T. Afzal, "Enhancing author assessment: An advanced modified recursive elimination technique (MRET) for ranking key parameters and conducting statistical analysis of top-ranked parameter," *International Journal of Data Science and Analytics*, pp. 1–23 (2024).

[39] R. Pates, P. Candon, K. Stenius, M. Miovský, J. O'Reilly, and T. Babor, *Publishing Addiction Science: A Guide for the Perplexed* (2017). https://doi.org/10.5334/bbd

[40] B. Ahmed, L. Wang, W. Hussain, G. Mustafa, and M. T. Afzal, "Investigating scholarly indices and their contribution to recognition patterns among awarded and non-awarded researchers," *International Journal of Data Science and Analytics*, pp. 1–18 (2025).

[41] B. Jin, L. Liang, R. Rousseau, and L. Egghe, "The r- and ar-indices: Complementing the h-index," *Chinese Science Bulletin*, vol. 52, no. 6, pp. 855–863 (2007). https://doi.org/10.1007/s11434-007-0145-9

[42] G. Mustafa, A. Rauf, and M. Tanvir Afzal, "MRET: Modified recursive elimination technique for ranking author assessment parameters," *PLoS One*, vol. 19, no. 6, e0303105 (2024).

[43] S. M. Adnan, W. Ahmad, I. Mahmood, G. Mustafa, and V. Dattana, "Enhancing text mining efficiency using an effective topic modeling approach," *Technical Journal*, vol. 29, no. 01, pp. 39–46 (2024).

[44] F. Ye and R. Rousseau, "Probing the h-core: An investigation of the tail core ratio for rank distributions," *Scientometrics*, vol. 84, no. 2, pp. 431–439 (2010). https://doi.org/10.1007/s11192-009-0099-6

[45] S. Alonso, F. J. Cabrerizo, E. Herrera-Viedma, and F. Herrera, "H-index: A review focused on its variants, computation and standardization for different scientific fields," *Journal of Informetrics*, vol. 3, no. 4, pp. 273–289 (2009). https://doi.org/10.1016/j.joi.2009.04.001

[46] W. Cheng, D. Zheng, X. Zheng, and H. Ni, "Combining referenced publication year spectroscopy and topic clustering to identify key knowledge foundations in scientometrics: An analysis of recipients of the Price Award," *Scientometrics*, pp. 1–23 (2025).

[47] A. Kanwal, N. Masood, G. Mustafa, M. A. Ghafoor, and S. Ayaz, "MK-SMOTE and M-SMOTE: Enhanced techniques for handling class imbalance problem," *Iran Journal of Computer Science*, pp. 1–19 (2025).

[48] H. Z. Mukhtiar, G. Mustafa, and M. T. Afzal, "Enhancing researcher evaluation in computer science: A novel index for impact assessment," *Knowledge and Information Systems*, pp. 1–18 (2025).

[49] G. Mustafa, M. T. Afzal, A. Rauf, and M. A. Khan, "Beyond publication numbers: A novel approach to academic ranking using evolutionary programming," *Evolutionary Intelligence*, vol. 18, no. 3, p. 62 (2025).

[50] M. Ameer and M. T. Afzal, "Evaluation of h-index and its qualitative and quantitative variants in neuroscience," *Scientometrics*, vol. 121, no. 2, pp. 653–673 (2019). https://doi.org/10.1007/s11192-019-03209-6

[51] Q.-u. Ain, H. Riaz, and M. T. Afzal, "Evaluation of h-index and its citation intensity based variants in the field of mathematics," *Scientometrics*, vol. 119, pp. 187–211 (2019). https://doi.org/10.1007/s11192-019-03009-y

[52] S. Ayaz and N. Masood, "Comparison of researchers' impact indices," *PLoS One*, vol. 15, no. 5, e0233765 (2020).

[53] A.-W. K. Harzing and R. Wal, "Google Scholar as a new source for citation analysis," *Ethics in Science and Environmental Politics*, vol. 8, no. 1, pp. 61–73 (2008). https://doi.org/10.3354/esep00076

[54] P. R. Kannari, N. S. Chowdary, and R. L. Biradar, "An anomaly-based intrusion detection system using recursive feature elimination technique for improved attack detection," *Theoretical Computer Science*, vol. 931, pp. 56–64 (2022). https://doi.org/10.1016/j.tcs.2022.07.030

[55] A. Rahangdale and S. Raut, "Machine learning methods for ranking," *International Journal of Software Engineering and Knowledge Engineering*, vol. 29, no. 06, pp. 729–761 (2019). https://doi.org/10.1142/S021819401930001X

[56] J. Zhang, C. Li, Y. Yin, J. Zhang, and M. Grzegorzek, "Applications of artificial neural networks in microorganism image analysis: A comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer," *Artificial Intelligence Review*, vol. 56, no. 2, pp. 1013–1070 (2023). https://doi.org/10.1007/s10462-022-10192-7

[57] D. H. Shirazi and H. Toosi, "Deep multilayer perceptron neural network for the prediction of Iranian dam project delay risks," *Journal of Construction Engineering and Management*, vol. 149, no. 4, 04023011 (2023). https://doi.org/10.1061/JCEMD4.COENG-12367

[58] M. Bilal, H. Israr, M. Shahid, and A. Khan, "Sentiment classification of Roman-Urdu opinions using Naive Bayesian, decision tree and KNN classification techniques," *Journal of King Saud University – Computer and Information Sciences*, vol. 28, no. 3, pp. 330–344 (2016). https://doi.org/10.1016/j.jksuci.2015.11.003

[59] A. R. Panhalkar and D. D. Doye, "Optimization of decision trees using modified African Buffalo algorithm," *Journal of King Saud University – Computer and Information Sciences*, vol. 34, no. 8, pp. 4763–4772 (2022). https://doi.org/10.1016/j.jksuci.2021.01.011

[60] A. Sinatra, D. Wang, P. Deville, C. Song, and A.-L. Barabási, "Quantifying the evolution of individual scientific impact," Science, vol. 354, no. 6312, p. aaf5239 (2016).