

Performance Evaluation of the Plagiarism Systems A Systematic Review

Nawzt H. Ali Al-Qazan, Shayma Mustafa Al- Mashhadany

Department of Computer Science, College of Computer Science and Mathematics, University of Mosul, Nineveh, Iraq
ha85dost@gmail.com, shaymamustafa@uomosul.edu.iq

Paraphrase plagiarism occurs when text is deliberately obfuscated to evade detection, deliberate alteration increases the complexity of plagiarism and the difficulty in detecting paraphrase plagiarism. Plagiarism systems play an important role in educational and research institutions. In the field of education, the use of these systems has increased in universities to help researchers select new research, which is why it has been important to ensure that universities have effective citation screening systems. Different metrics were used to evaluate the performance of these systems. All of these metrics affect the performance and usability of the impersonation system. This paper presents a systematic literature review (SLR) to obtain an overview of the existing studies of distinct criteria for evaluating the performance of textual plagiarism systems and each criterion is briefly explained to give the basic idea behind it. Overall, the aim of this paper is to provide a modern introduction and short review of these criteria. A systematic literature review is performed according to the guidelines for conducting SLR in software engineering. The study relied on a comprehensive set of research papers collected from electronic libraries published during a time period from 2013 to 2023. The results of this SLR include information about metrics for evaluating the performance of plagiarism systems according to previous research. During a specific period it can assist researchers in this field by providing an overview of current researches in this field. Moreover, it may serve as a first step towards a great SLR-assisted explanation of the subject.

Keywords: Plagiarism systems plagiarism detection, Performance Evaluation, Performance Indicators



This is an open access article under the [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)

1. INTRODUCTION

Plagiarism is defined as “using words, ideas, or work products attributable to another identifiable person or source without attributing the work to the source from which it was obtained” [1]. A number of plagiarism systems are designed every day. These systems will not have the same degree of performance and efficiency. If the plagiarism system performs poorly, the researcher simply leaves this system and uses another plagiarism system, and there is no chance of bringing this researcher back to using a poorly performing system again [2] - [6]. Therefore, in order to improve the performance and efficiency of these systems, it is important to design impersonation systems with some new characteristics, such as: Accessibility, Accuracy of Results, Clarity Report, Diversity Of Text Format, Download Time, Expandability, Open Source, Plagiarism Security, Supported Languages, and Validation Time [7] - [10]. This can be done by defining good system standards, as the performance of the plagiarism system depends on some measurable standards that provide an effective way to develop this system. However, the performance evaluation process has become a challenge with entirely new systems [11]. In general, the process of evaluating the performance of plagiarism systems has become a fundamentally valuable topic and is subject to continuous development, especially in the field of performance metrics of plagiarism systems [12]. Many metrics have been developed to measure the performance of impersonation systems as shown in Table 1. These indicators obtained high percentages that make them usable in designing a plagiarism system.

This paper presents a systematic literature review to provide an overview of primary studies on evaluating the performance of plagiarism detection systems since 2013. The motivation is to identify the available evidence on this topic and identify a research gap in evaluating appropriate performance criteria. After the introduction in Section 1, the structure of this paper is as follows. Section 2 includes the framework of the systematic review of the literature, and Section 3 contains a discussion of the research questions that explain the performance measures of the plagiarism detection system, conducting statistics, issuing the appropriate weight for each indicator, and then dividing the performance indicators into main and secondary according to their weight, while the conclusions are presented in Section 4.

Table.1. Plagiarism systems Performance Indicators

No.	Indicators	Definition
1	Accessibility	It measures the easy of use and accessibility of the plagiarism detection system (Farah al., 2021) .
2	Accuracy of Results	It measures the accuracy of the verification results of the plagiarism detection system, and this indicator depends mainly on the size of the database that the plagiarism detection system possesses (Hakkun et al. , 2018) .
3	Clarity Report	It measures the clarity of the report to the user resulting from the verification process in terms of the use of easy language, clear values, and a clear interface for the report in general (Uman,2021) .
4	Diversity Of Text Format	It measures the diversity of text formatting in plagiarism detection systems. This indicator must be the highest possible value to achieve high reliability (Mohamed et al. , 2023).
5	Download Time	It measures the time of uploading the file whose attribution is to be measured to the plagiarism detection system and is measured in seconds (Prashanth & Subramanya ,2023).
6	Expandability	It measures the possibility of expanding the database of plagiarism detection systems to include all recent research in real time (Mehdi et al. , 2022).
7	Open Source	It measures plagiarism detection systems in terms of cost. Is the system free, not free, or limited?(Marko,2022)
8	Plagiarism Security	The security index of plagiarism detection systems measures the possibility of unauthorized persons accessing files in the system database(Ahmed et al. , 2012).
9	Supported Languages	The language support index measures the average number of languages supported by the plagiarism detection system (Mohran et al. , 2018).
10	Validation Time	This indicator measures the time when the plagiarism detection system verifies the file and sends it to the researcher (Ivana et al. , 2012).

2. METHOD

The process of selecting and classifying research from as much current and previous literature as possible that is relevant to an issue is known as a systematic literature review. When applied to a particular topic, it often produces a summary and map of its findings by classifying different types of research reports along several different dimensions [13] - [17]. These investigations are primarily proposed for research areas with very broad topics and little relevant information discovered during initial field studies of research. When conducting a detailed review, the only aims are to identify evidence relevant to the research questions and identify any knowledge gaps that could inform future study [18]. In this study, a systematic study of performance indicators of plagiarism detection systems was conducted, as it appears to be a broad topic with different research focus areas. However, no existing research has conducted a systematic review of the literature in this area. This section describes the review protocol that includes the basic process of defining the research questions, defining the search strategy, selecting relevant literature, and the methodological map for evaluation.

2.1. Selected Primary Studies

This paper presents a systematic review of the work done in the field of plagiarism detection systems (PDS) performance indicators and in order to get a broad view, various papers and journals have been searched and selected the publications that related to this study within the time span 2013 to 2023. After selecting the publications related to the study within this period, 30 articles have been found that very closely to the plagiarism systems performance indicators, The search strings that Table (2) are found among academic databases to locate articles having these strings in their abstracts, titles, and keywords. Accordingly, famous online academic databases like ACM, IEEE, Science Direct, Springer, Google Scholar, Taylor & Francis, and Wiley are used.

2.2. Research Questions

The aim of this study is to obtain a presentation and overview of the current and previous research in the field for evaluating the performance indicators of plagiarism detection systems (PDS) through the following questions. The overall objective is defined in these research questions.

Research Question 1: What are the performance indicators of plagiarism detection systems that previous researches focused on?

Research Question 2 : What are The Plagiarism Detection Systems Being Studied in This Researches?

Research Question 3: What are the plagiarism systems Performance Indicators?

2.2.1 Research Question 1

What are The Different Indicators Used in Evaluation The Performance of plagiarism systems?

The indicators of evaluating the performance of (PDS) were presented in Table 1, these indicators were collected after a comprehensive study of the previous researches stated in table (2), that presents a group of previous research specialized in evaluating the performance of (PDS) within the period from 2013 to 2023 where the metrics used in each research were determined .

The Utilization weight of the previous studies has been calculated for each Indicator, based on the percentages gained by the evaluation indicator from prior studies. The number of studies for each indicator and the percentages of these Indicator are displayed in Table (3).

Table .2. plagiarism detection systems Performance Indicators Based On The Previous Researches

No	Researchers	Accuracy of Results	Clarity Report	Diversity Of Text Format	Download Time	Expandability	Open Source	Plagiarism Security B	Supported Languages	Validation Time
1	(Mishra & Sudhakar, 2023)	✓	✓	✓	✗	✓	✓	✓	✗	✓
2	(Victor , 2023)	✓	✗	✗	✓	✓	✓	✗	✓	✓
3	(Yetthapu & Sudheer, 2023)	✗	✗	✓	✗	✓	✓	✗	✓	✗
4	(Ayoub , 2023)	✗	✓	✗	✓	✓	✓	✗	✓	✓
5	(Michael , 2022)	✓	✗	✓	✗	✓	✗	✓	✓	✓
6	(Keerthana et al. , 2022)	✓	✓	✗	✓	✓	✓	✓	✓	✓
7	(Mehdi et al. , 2022)	✓	✗	✓	✓	✓	✓	✗	✓	✗
8	(Jambi et al. , 2022)	✓	✗	✗	✓	✗	✗	✓	✗	✓
9	(Mahshad , 2022)	✗	✓	✓	✗	✓	✓	✗	✓	✓
10	(Hussain , 2021)	✓	✓	✓	✓	✗	✗	✓	✓	✓
11	(Farah et al., 2021)	✗	✗	✗	✓	✓	✓	✗	✗	✓
12	(Siwar et al. , 2021)	✓	✓	✗	✗	✓	✓	✗	✗	✓
13	(Cynthia , 2020)	✗	✓	✓	✗	✗	✗	✓	✗	✗
14	(Tomas et al., 2020)	✓	✓	✗	✓	✗	✓	✓	✓	✗
15	(Alsallal et al., 2019)	✗	✓	✓	✗	✓	✓	✗	✗	✓
16	(Taylor & Francis, 2019)	✗	✗	✗	✓	✓	✗	✗	✓	✓
17	(Francis et al., 2019)	✗	✗	✗	✓	✗	✗	✗	✓	✗
18	(Konstanz , 2019)	✗	✓	✓	✗	✓	✓	✗	✓	✓
19	(Olfat ,2018)	✗	✓	✗	✗	✗	✓	✗	✓	✓
20	(Hakkun et al. , 2018)	✗	✓	✓	✗	✓	✗	✗	✗	✗
21	(Anton & Marina , 2018)	✗	✓	✗	✓	✗	✗	✗	✗	✓
22	(Yousef , 2017)	✗	✓	✗	✗	✓	✓	✗	✓	✓
23	(Marius , 2017)	✗	✓	✗	✓	✗	✗	✓	✓	✗
24	(Fredrik , 2017)	✗	✓	✗	✗	✓	✓	✗	✓	✓
25	(Curtis & Vardanega , 2016)	✗	✓	✗	✗	✗	✗	✓	✓	✗
26	(Asad et al., 2015)	✗	✓	✗	✗	✗	✗	✗	✓	✗
27	(Tuomo & Maxim, 2014)	✗	✓	✓	✗	✗	✓	✗	✓	✓
28	(Angelos , 2013)	✗	✓	✗	✗	✗	✓	✓	✓	✓
29	(Ahmed et al. , 2013)	✓	✓	✗	✓	✓	✓	✓	✓	✓
30	(Bensalem, 2013)	✗	✓	✗	✓	✗	✓	✓	✓	✓
Sum of Researches Used for Indicators		10	22	11	14	17	19	8	23	20

Table .3. Indicators and References in PDS

No	Indicators	References	Utilization rate
1	Accessibility	(1,2,5-8,10,12,14,29)	33.33 %
2	Accuracy of Results	(1,4,6,9,10,12-15,18-30)	73.33 %
3	Clarity Report	(1,3,5,7,9,10,13,15,18,20,27)	36.66 %
4	Diversity Of Text Format	(2,4,6-8,10,11,14,16,17,21,23,29,30)	46.66 %
5	Download Time	(1-7,9,11,12,15,16,18,20,22,24,29)	56.66 %
6	Expandability	(1-4,6,7,9,11,12,14,15,18,19,22,24,27-30)	63.33 %
7	Open Source	(1,5,6,8,10,14,23,29)	26.66 %
8	Plagiarism Security	(2-4,6,8-10,12-29,22-25,27-30)	76.66 %
9	Supported Languages	(1-7,9,10,14,16,17,26)	43.33 %
10	Validation Time	(2,4-6,8-12,15,16,18,19,21,22,24,27-30)	66.66 %

These Indicators Could Be Characterized As Follows :

- *Accessibility* – It measures the user's ability to access the plagiarism detection system easily and clearly (Farah al., 2021). Now it's expected usage rate is (33.33%), as shown in Table (3).
- *Accuracy of Results* - It measures the accuracy of the information after verification by the plagiarism detection system (Hakkun et al. , 2018). Now it's expected usage rate is (73.33%), as shown in Table (3).
- *Clarity Report* - It measures the clarity of the information resulting from the verification process of the plagiarism detection system (Uman,2021). Now it's expected usage rate is (36..66%), as shown in Table (3).
- *Diversity Of Text Format* - It measures the diversity of text formatting in plagiarism detection systems(Mohamed et al. , 2023). Now it's expected usage rate is (46..66%), as shown in Table (3).
- *Download Time* – It measures the time the file is uploaded to the plagiarism detection system (Prashanth & Subramanya ,2023).Now it's expected usage rate is (56..66%), as shown in Table (3).
- *Expandability* - It measures the possibility of expanding the database of plagiarism detection systems(Mehdi et al. , 2022) .Now it's expected usage rate is (63..33%), as shown in Table (3).
- *Open Source* – It measures plagiarism detection systems in terms of cost. Is the system free, not free, or limited?(Marko,2022). Now it's expected usage rate is (26.66%), as shown in Table (3).
- *Plagiarism Security* – It measures the possibility of unauthorized persons accessing files in the system database(Ahmed et al. , 2012).Now it's expected usage rate is (76.66%), as shown in Table (3).
- *Supported Languages* – It measures average number of languages supported by the plagiarism detection system (Mohran et al. , 2018).Now it's expected usage rate is (43.33%), as shown in Table (3).
- *Validation Time* – It measures the time when the plagiarism detection system verifies the file and sends it to the researcher (Ivana et al. , 2012).Now it's expected usage rate is (66.66%), as shown in Table (3).

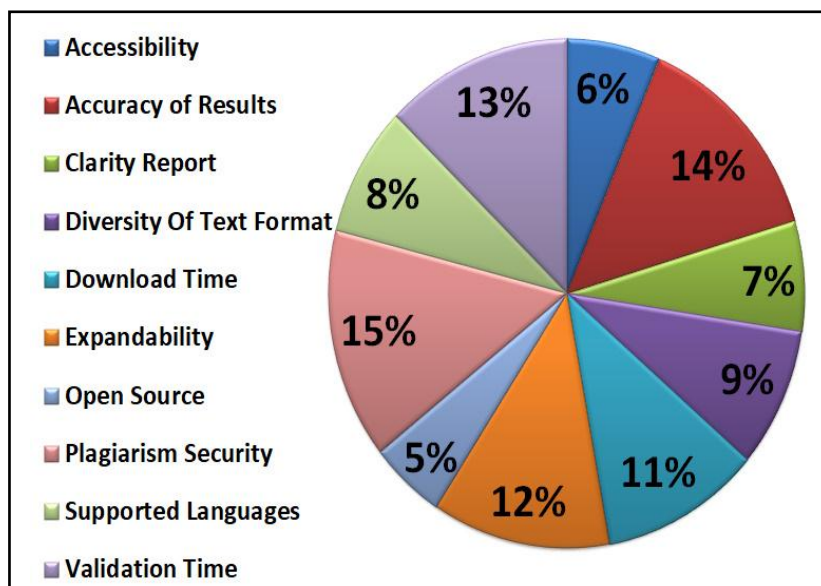


Fig.1. PDS performance Indicators Percentages Based On Previous Studies

2.2.2 Research Question 2

What are the plagiarism detection systems being studied in this research?

Plagiarism detection systems or intellectual theft is defined as the claim by one person, explicitly or implicitly, of writing what another has written or copying what others have written, in whole or in part, without proper excuse or acknowledgment; In short, it is false or gives the impression that you wrote what someone else wrote. It is considered a fraudulent act. The most popular and used plagiarism detection systems by researchers were chosen for the purpose of these study, which is evaluating the performance of these systems, and the number of these systems are 8, arranged alphabetically, as shown in Table 4.

Table .4. The Utilization Rate of each Indicators in the References

No	Plagiarism Detection Systems	URL
1	Al-Sabt Center for Scholars	https://network.srp-center.iq
2	Duplicate Checker	http://www.duplichecker.com
3	Small Seo Tools Plagiarism Checker	http://www.smallseotools.com/plagiarism-checker
4	Turnitin	https://www.turnitin.com
5	The Plagiarism Checker	http://www.dustball.com/cs/plagiarism.checker
6	Plagiarisma.Net	http://www.plagiarisma.net
7	Plag Tracker	http://www.plagtracker.com
8	Viper	http://www.scanmyessay.com

2.2.3 Research Question 3

What are The Plagiarism Detection System Performance Indicators?

This focuses on the performance indicators of plagiarism detection systems, which are important elements for evaluating the performance of these systems. The section also explains the extent to which these systems pay attention to performance indicators, as the eight plagiarism detection systems were chosen for the purpose of the study, as shown in Table 4. These indicators are divided into primary and secondary indicators, that can be measured quantitatively and qualitatively. The indicators for evaluating the performance of different plagiarism detection systems will be explained in order to choose the two most commonly used main evaluation indicators to be used in the design phase of the proposed tool that will be designed to evaluate the performance of plagiarism detection systems.

3. RESULTS AND DISCUSSION

Plagiarism detection system performance indicators are the basic components that affect the quality of any plagiarism detection system. These indicators must be evaluated to verify the performance and effectiveness of plagiarism detection. These indicators are divided into the main performance indicators of the plagiarism detection system Key Performance Indicators (PDS - KPIs) that effectively contribute to calculating the performance evaluation of these systems, and the secondary performance indicators of the plagiarism detection system Secondary performance indicators (PDS - SPIs) that measure other aspects that will be explained in this section, and figure (3) shows the main and secondary performance indicators of plagiarism detection systems.

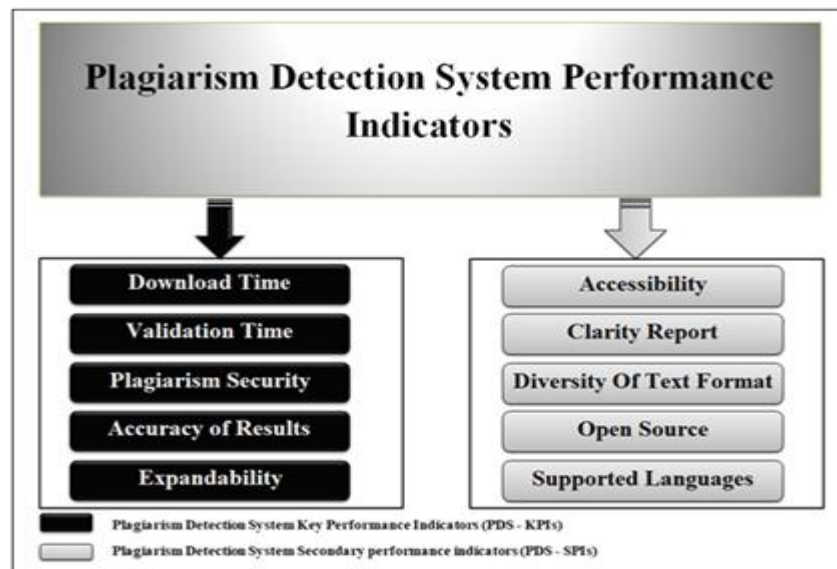


Figure 3 - The Performance Plagiarism Detection System Indicators Types

3.1 The Main Performance Indicators of Plagiarism Detection System (PDS - KPIs)

- Download Time (DT)

Loading time is one of the most important key indicators of the performance of any software application, and the time of loading files in plagiarism detection systems is an important factor for the researcher so that the researcher can make modifications in a timely manner and then re-check the plagiarism. The Turnitin plagiarism detection system is considered faster than the AL-Sabt plagiarism detection system and the rest of the systems mentioned in this research in downloading files and thus has a shorter download time (Prashanth and Subramanya, 2023).

- Validation Time (VT)

Verification time is also another important measure. Some plagiarism detection systems take a lot of validation time but also need a little time. The more papers to compare, the more validation time to use, and the more sources to compare, the longer the validation time. Among the eight programs studied. Turnitin plagiarism detection system give the fastest result as every sentence is checked instantly without delay. While Plagiarism Checker has produced a quick result for the first document but for subsequent documents there will be a long delay until the result is produced, as the status is "In Progress" and during this status no further document can be verified. For the other six programs (Plag Tracker ,Viper, AL-Sabt plagiarism detection system, Small Seo Tools Plagiarism Checker, Plagiarisma.Net, and Dupli Checker) they all require almost identical time which can be related to the number of words in the document; But Plagiarisma Net and Dupli Checker received a low score in this test because both programs did not display stage progress, so users have no information about how long the process will take (Ivana et al., 2012).

- Plagiarism Security (PS)

This metric means security of documents in validation. Some problems arise in that the free PDS sometimes takes scanned documents to be saved in its databases, this is a problem for some people. Among the official websites of these eight plagiarism detection systems, Turnitin plagiarism detection system,Plagiarisma.Net and Plag Tracker stated clearly that they did not save any user-uploaded content. While Viper states that it saves all scanned articles within a secure database, its goal is to get an accurate plagiarism report from previously submitted articles. According to Viper, no one else can access the database including lecturers, universities and other plagiarism checkers (Viper 2012). After eight months of examining the files, Viper uploads the article to its database and it appears on one of our study sites, so that other students can use it to help them in writing their own essays (Viper 2012). As for other plagiarism detection systems, Le. The Plagiarism Checker, AL-Sabt plagiarism detection system ,Duplicate Checker, and Small Seo Tools Plagiarism Checker there is no information about this matter on their official websites (Ahmed et al., 2012).

- Accuracy of Results (AOR)

Accuracy is the most important thing in plagiarism detection systems. To evaluate the accuracy of each program the test is based on the sample article provided. As mentioned earlier, each article sample contains some percentage of plagiarism taken from different sources (Haakon et al., 2018). A program with good accuracy is one that returns the amount of percentages that are close to the plagiarism percentages, and the sources are also displayed according to the sources mentioned in the sample article. But not every plagiarism detection system gives the result of possible plagiarism rates. Turnitin plagiarism detection system ,The Plagiarism Checker ,AL-Sabt plagiarism detection system ,and Duplicate Checker showed only suspected sources as cited sources (Amalia, 2013).

- Expandability

The expandability of the database of plagiarism detection systems is an indicator that ensures that research is examined with high accuracy, but the expansion must take place in real time to publish research in scientific databases with the ability to update. This guarantees the highest accuracy of these systems (Mehdi et al. , 2022).

3.2 The Secondary Performance Indicators Of Plagiarism Detection System (Pds - Spis)

- Accessibility

Accessibility is a term that refers to the ability of accessing devices or systems easily in order to achieve a specific goal (Farah al., 2021). Accessibility also means the ways to measure the ease of use and study the principles behind the efficiency or elegance of those devices or systems. These parameters are very subjective, and a high or low rating for this parameter depends on the needs of the user. Some PDSs require a user account to be installed and registered to obtain software services. For users who only occasionally need anti-spoofing software, the installation and registration steps will take some time. But for users who regularly use the anti-plagiarism service such as teachers, lecturers, librarians, etc., this is not a problem, because once the installation and account registration is completed, we simply select the file or files you want to check for plagiarism and the program will generate the result. Among these eight programs, only Viper needs to be installed. This program is small in size, measuring only 1 MB, and can be downloaded from the official Viper websites. Besides having to install the program, the user must also have a Viper account to run this service. In addition, this parameter also evaluates

the methods for entering the article into plagiarism checks. There are several options for how to scan documents that are commonly used by the Plagiarism Detection System (PDS), such as copying and pasting plain text into a text box and uploading a text file. Among the eight programs that provide these two features is Plagiarisma. Net and Duplicate Checker While these two features are available in the Plagiarism Checker and Plagiarism Tracker, but for free account users, they can only scan documents by copying and pasting plain text into the text box. For Small Seo Tools Plagiarism Checker, only one feature is offered. Copy and paste the plain text into the text box. As for the Viper feature that can be used simply by uploading the document (Farah al., 2021).

- Clarity Report (CR)

The Clarity Report is indication of the validation result. A good report is one that the users can easily understand. In this study, among the eight programs, only Al-Sibt plagiarism detection system and Turnitin plagiarism detection system feature a clear report and also Plagiarisma.Net and Viper are able to generate the report in file format. As for the PDS that showed matching content or originality in percentage, Viper, Plagiarisma.Net, Small Seo Tools Plagiarism Checker, and the Plagiarism Checker, while Dupli Checker and The Plagiarism Checker only showed link content/sites that had similarity to the verified article. To see the similarity of text, users must open the sites, but unfortunately they do not highlight the similarity between them. PDS reports that highlight text potentially plagiarism are Viper, Plag Tracker, and Plagiarisma.Net. For a PDS in which a report is not generated in file format, the resulting verification view is an html file that can be saved and reopened again offline, with the exception of the plagiarism checker. The report page of the software still needs to be connected to the internet to display the result of the plagiarism scan (Oman, 2021).

- Diversity Of Text Format (DOTF)

Diversity of verifiable text format is a deciding factor for some people. Since in evaluating PDS it is free PDS, sometimes there are restrictions imposed by the software development company. Among these eight programs, only Turnitin plagiarism detection system ,Plagiarisma.Net, Viper, AL-Sabt plagiarism detection system, and Duplicate Checker are able to validate documents in different text format. Plagiarisma.Net is able to validate documents in most types of text formats, e.g. pdf, doc, docx, rtf, odt, txt and html. Viper can validate your the document using doc, docx, txt, rtf and pdf text format. While Dupli Checker is only able to validate documents in docx and txt text format. To get a free account with Plag Tracker version. Plagiarism Checker and Plagiarism Checker are not able to check article plagiarism using some text formats because the software validation method was done by cutting and pasting the plain text from the validated documents or article into a text box in the software menu (Muhammad et al, 2023).

- Open Source (OS)

This indicator gives three possibilities, which are a free, not free, and limited plagiarism detection system. An example of the plagiarism detection system specified in the Al-Sabt Center for Researchers. When two papers are submitted at the same time, only one paper will be verified and the second paper will be ignored (Marko,2022).

- Supported Languages

Plagiarism detection systems support a number of languages. The more languages, the more the plagiarism detection system is used, but this also increases the system's database. The database can be distributed on different distributed servers, and thus this distribution contributes for reducing download time and verification time (Mohran et al., 2018). The Turnitin plagiarism detection system was produced by iParadigms company, as the system only deals with texts and since . In 2008, the producing company announced the adoption of (32) languages, including English and Arabic, and the system can deal with them in the process of detecting copied content (Naseer ,2017).

3. CONCLUSION

This paper presented a systematic literature review to obtain an overview of the existing researches in the field of performance evaluation of text plagiarism detection systems, as the researchers has proven that up to 90% of university students needs the plagiarism systems, which affects academic integrity. As a result, it has become necessary to evaluate these systems that contribute of detecting plagiarism. According to the most used performance metrics, several major and secondary performance indicators, that are currently used to evaluate the performance of various plagiarism systems. In this research, the researches were chosen between the years 2013 and 2023 to conduct a complete study on the evaluation performance indicators, which are: accessibility, accuracy of results, clarity of report, text, download time, scalability, open source ,Spoofing system security, language support, and verification time. In this study, a statistical analysis in terms of the most, used performance indicators is presented. After a comprehensive study of previous works, it became clear that most researchers focused on the speed of downloading files, speed of verification, accuracy of results, security of plagiarism, and scalability in recent years, and they obtained high exploitation rates, which made them useful for system design. They were dealt with as key performance indicators, while the indicators obtained accessibility and clarity of the results report. Accessibility, clarity report, open source, diversity of text format, open source, and supported

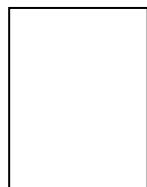
languages had a low usage rate and were treated as secondary performance indicators according to previous studies.

REFERENCES

- [1]. Ayoub Ali Mohammed Saeed,(2023). " Designing and Implementing Intelligent Textual Plagiarism Detection Models", Doctoral dissertation, College of Computer Science and Mathematics, University of Mosul, Nineveh, Iraq.
- [2]. Mohamed A. El-Rashidy, Ramy G. Mohamed,Nawal A. El-Fishawy,Marwa A. Shouman, (2023)." An effective text plagiarism detection system based on feature selection and SVM techniques". Multimedia Tools and Applications, Vol. 1, No. 2 ,PP: (1-39).
- [3]. Prashanth Kumar HM1 & Subramanya Bhat .(2023). "A Plagiarism Checker: Analysis of time and space complexity". International Research Journal of Engineering and Technology (IRJET). Volume: 10 Issue: 09,PP:(52-56).
- [4]. Mishra, Sudhakar,(2023)."Enhancing Plagiarism Detection: The Role of Artificial Intelligence in Upholding Academic Integrity" . Library Philosophy and Practice (e-journal). 7809. <https://digitalcommons.unl.edu/libphilprac/7809>
- [5]. Victor Ughakepen Thompson,(2023)." Integrating State-of-the-art NLP Tools into Existing Methods to Address Current Challenges in Plagiarism Detection", Doctoral dissertation, University of Sunderland.
- [6]. Yetthapu, Sudheer, (2023). "Source Code Plagiarism Detection Using JPlag & Stack Overflow Data" ,Masters Theses & Specialist Projects. Paper 3620. <https://digitalcommons.wku.edu/theses/3620>
- [7]. Michael Owen Harris,(2022)." Differences Of Attitudes And Norms Toward Plagiarism Between First- And Continuing-Generation College Students ", Doctoral dissertation, Liberty University .
- [8]. Keerthana T V , Pushti Dixit , Rhuthu Hegde , Sonali S K , and Prameetha Pai,(2022). "A Literature Review on Plagiarism Detection in Computer Programming Assignments" , International Research Journal of Engineering and Technology (IRJET), Volume: 09 Issue: 03, PP : (1073- 1083).
- [9]. Mehdi Abdelhamid1, Sofiane Batata1 and Faiçal Azouaou1,(2022). "A Survey of Plagiarism Detection Systems: Case of Use with English, French and Arabic Languages", Open Access Giving Week,val(1),no(10), PP : (1- 28).
- [10]. Jambi, K.M.; Khan, I.H.; Siddiqui, M.A.,(2022). "Evaluation of Different Plagiarism Detection Methods: A Fuzzy MCDM", Appl. Sci., Volume: 01 Issue: 01, PP : (1- 15).
- [11]. Mahshad,Davoodifard,(2022). " Automatic Detection of Plagiarism in Writing", Studies in Applied Linguistics & TESOL at Teachers College, Columbia University, Vol. 21, No. 2, PP :(54 - 60) .
- [12]. Marko,(2022). "Textual and Source Code Plagiarism in Academic Environment: a Serbian perspective". Plagiarism detection conference 2022, University of Belgrade School of Electrical Engineering Department of Computer Engineering and Informatics Serbia.
- [13]. Farah Khaled, Mohammed Sabbih H. Al-Tamimi . (2021). "Plagiarism Detection Methods and Tools: An Overview". Iraqi Journal of Science, Vol. 62, No. 8, PP : (2771-2783).
- [14]. Uman,(2021). "Plagiarism Checker X Originality Report". International Review of Automatic Control (I.R.E.A.CO.), Vol. 10, N. 2 ,PP: (6019-6059).
- [15]. Hussain A. Chowdhury, (2021) . "Plagiarism: Taxonomy, Tools and Detection Techniques", Master's Thesis , Dept. of CSE, Tezpur University.
- [16]. Farah K. AL-Jibory, Mohammed S. H. Al- Tamimi, (2021) ." Hybrid System for Plagiarism Detection on A Scientific Paper" . Turkish Journal of Computer and Mathematics Education, Vol.12, No. 13 ,PP: (5707-5719)
- [17]. Siwar Nadhri, Maryam Elamine and Lamia Hadrich Belguith,(2021) . "Automatic evaluation of existing plagiarism detection tools". Tunisian Algerian Conference on Applied Computing (TACC 2021), December 18–20, 2021, Tabarka, Tunisia.
- [18]. Cynthia D. Scott , (2020) . "The Plagiarism-Proof Policy Handbook: A Multidimensional-Systems Approach To Foster Student Writing Accountability and Prevent Plagiarism" . Master's Thesis, Faculty of The Liberty University Department of English .
- [19]. Tomas Foltýnek, Dita Dlabolova1, Alla Anohina-Naumeca, Salim Razi, Július Kravjar, Laima Kamzola, Jean Guerrero-Dib, Ozgür Celik & Debora Weber-Wulff , (2020) . "Testing of support tools for plagiarism detection" . International Journal of Educational Technology in Higher Education, Vol.17, No. 46 ,PP: (1-32) .
- [20]. Alsallal, M, Iqbal, R, Palade, V, Amin, S & Chang, V (2019) . "An integrated approach for intrinsic plagiarism detection" , Future Generation Computer Systems, vol. 96, No. 1 ,PP: (700-712) . <https://dx.doi.org/10.1016/j.future.2017.11.023>.
- [21]. Taylor & Francis , (2019) . "Design and Evaluation of Plagiarism Prevention and Detection Techniques in Engineering Education" . Higher Education Pedagogies, Vol. 4, No. 1, PP : (197–208) .
- [22]. Francis & Taylor , (2019) . "Design and Evaluation of Plagiarism Prevention and Detection Techniques in Engineering Education" . Higher Education Pedagogies, Vol. 4, No. 1, PP : (197–208) .
- [23]. Konstanz, Germany , (2019) . "Academic Plagiarism Detection: A Systematic Literature Review", ACM Computing Surveys, Vol. 52, No. 6, PP : (1– 46) .
- [24]. Lame, G. (2019). "Systematic Literature Reviews: An Introduction" , in Proceedings of the 22nd International Conference on Engineering Design (ICED19), Delft, The Netherlands, 5-8 August 2019. DOI:10.1017/ dsj.2019.169 .
- [25]. Olfat Meraj Mirza, (2018) , "Style Analysis for Source Code Plagiarism Detection", Doctoral dissertation, Department of Computer Science, Warwick University.
- [26]. Hakkun Elmunsyah , Hary Suswanto, Khoirudin Asfani , & Wahyu Nur Hidayat ,(2018) ." The Effectiveness of Plagiarism Checker Implementation in Scientific Writing for Vocational High School ". Advances in Social Science,

- Education and Humanities Research (ASSEHR), volume 201 ,International Conference on Indonesian Technical Vocational Education and Association (APTEKINDO 2018) , PP: (192-196) .
- [27]. Anton Belyy, Marina Dubova,(2018). “Improved Evaluation Framework for Complex Plagiarism Detection” . Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Short Papers), pages:(157–162).
- [28]. Mohran H. Al-Bayed, Samy S. Abu-Naser, (2018). “Intelligent Multi-Language Plagiarism Detection System”. International Journal of Academic Information Systems Research (IJAIRS), Vol. 2 ,Issue 3, PP: (19-34).
- [29]. Yousef k. Sanjalawe & Mohammed Anbar , (2017) . “The Plagiarism Detection Systems For Higher Education- A Case Study In Saudi Universities” . International Journal of Software Engineering & Applications (IJSEA), Vol.8, No.2, PP: (33-49) .
- [30]. Naseer Ali Husieen ,(2017). "Getting to know plagiarism detection programs (scientific theft of scientific research)". Proceedings of the third annual scientific conference of the College of Arts/University of Wasit, Held under the slogan (Publishing Movement and Development of Human Studies) for the period November 30 - December 1, 2017 AD,PP: (93-110).
- [31]. Naseer Ali Husieen ,(2017). "Getting to know plagiarism detection programs (scientific theft of scientific research)". Proceedings of the third annual scientific conference of the College of Arts/University of Wasit, Held under the slogan (Publishing Movement and Development of Human Studies) for the period November 30 - December 1, 2017 AD,PP: (93-110).
- [32]. Naseer Ali Husieen ,(2017). "Getting to know plagiarism detection programs (scientific theft of scientific research)". Proceedings of the third annual scientific conference of the College of Arts/University of Wasit, Held under the slogan (Publishing Movement and Development of Human Studies) for the period November 30 - December 1, 2017 AD,PP: (93-110).
- [33]. Marius Popescu, Csiki Laszlo, and Antoanela Naaji, (2017) . “Software for Assessing the Performance of Anti-plagiarism Programs”. MATEC Web of Conferences 125, 04005 (2017),PP: (1-7).
- [34]. Fredrik Albrektsson, (2017) . “Detecting Sockpuppets in Social Media with Plagiarism Detection Algorithms” . Master’s Thesis, KTH University, Stockholm, Sweden.
- [35]. Curtis, G. J., & Vardanega, L. (2016). “Is plagiarism changing over time? A 10-year time-lag study with three points of measurement” . Higher Education Research & Development, Vol. 35, No. 6 , PP: (1167–1179) . <https://doi.org/10.1080/07294360.2016.1161602> .
- [36]. Asad Abdia, Norisma Idris , Rasim M. Alguliyev, & Ramiz M. Aliguliyev, (2015) .” PDLK: Plagiarism detection using linguistic knowledge”. Expert Systems With Applications 42 (2015), PP:(8936–8946).
- [37]. Tuomo Kakkonen & Maxim Mozgovoy ,(2014) . “An Evaluation of Web Plagiarism Detection Systems for Student Essays” .Department of Computer Science and Statistics, University of Joensuu, Finland, Conference Paper,PP: (99-103).
- [38]. Angelos Konstantinidis , (2013) . “Plagiarism : Examination of Conceptual Issues and Evaluation of Research Findings on Using Detection Services” . Contemporary Educational Technology, Vol. 4, No. 3 ,PP : (212-221).
- [39]. Amalia, (2013) . “Performance evaluation of free anti-plagiarism software” .Annual International Conference Sylah KualaUnlversity (AIC Unsyiah) 2013 , PP: (30-36).
- [40]. Bensalem, I., Rosso, P.,and Chikhi, S. (2013). “A new corpus for the evaluation of Arabic intrinsic plagiarism detection”. En Information Access Evaluation. Multilinguality, Multimodality, and Visualization. Springer Verlag (Germany). PP: (53-58).
- [41]. Ahmed Hamza Osman, Naomie Salim, and Albaraa Abuobieda, (2013). “Survey of Text Plagiarism Detection” Computer Engineering and Applications ,Vol. 1, No. 1, PP: (37-45).
- [42]. Ivana Bosnic, Branko Mihaljevic, Marin Orlic and Mario Zagar, (2012) .” Source Code Validation And Plagiarism Detection”, CSEDU2012-4thInternationalConferenceonComputerSupportedEducation,PP: (149-154).
- [43]. Fishman, T. (2009). “We know it when we see it is not good enough: Toward a standard definition of plagiarism that transcends theft, fraud, and copyright”. Proceedings of 4th Asia Pacific Conference on Educational Integrity (4APCEI) PP: (1-5)

BIOGRAPHIES OF AUTHORS



Nawzt H. Ali Al-Qazan, A master’s graduate student at the University of Mosul, College of Computer Science and Mathematics/Computers Department. I have been working as a computer teacher for more than ten consecutive years. I finished my Bachelor’s degree in the College of Education in Computer Science in 2009 with a high grade of good. I have a TOT training certificate recognized by the German Society and have worked for several years in media field