وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم علوم الحاسوب

تقييم أداء أنظمة الاستلال النصي باستخدام مقاييس هندسة البرمجيات

رسالة مقدمة
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة ماجستير علوم في
علوم الحاسوب

من قبل
نوزت حسن علي حسن

بإشراف
م.د. شيماء مصطفى محي الدين ذنون

١٤٤٦ هـ ٢٠٢٤ م

الخلاصة

تعد أنظمة وبرامج كشف الانتحال من أهم التطبيقات البرمجية التي تخدم الباحثين والطلاب في مجال التعليم والحفاظ على النزاهة الأكاديمية، وبسبب ذلك يتم تطوير هذه البرامج وصيانتها باستمرار ومعالجة نقاط الضعف فيها للحصول على أعلى نسبة فائدة ممكنة من أداءها ومنع السبل التي يستخدمها من يمارسون الانتحال للتشكيك والتلاعب بنزاهة البحوث والدراسات التي يتم الإفادة منها في مجال التطور والارتقاء العلمي.

فضلاً عن ذلك نحتاج إلى دراسة وتقييم تلك البرامج للتعرف أكثر على جودة الأداء ومدى إمكانية رفع مستوى المعالجات التي تقدمها.

تم دراسة العديد من الأعمال السابقة في مجال تقييم أداء برامج كشف الانتحال والمؤشرات التي تدخل في تقييم أداء تلك البرامج وتحليل بعض تلك المؤشرات لمعرفة اهم الفروقات بين البرامج، وقد شملت هذه الدراسة جمع البحوث والأعمال في مجال التقييم من مجلات ومكتبات إلكترونية منشورة من العام 2013 لغاية العام 2023.

وقد تم تحديد نماذج المستندات في تخصصات مختلفة منها مجالات الصحة وعلوم الحاسوب والرياضة، ليتم بعدها تقسيم هذه المقالات والنصوص إلى أكثر من حالة، مثلاً حالة مستند يحتوي على مواد منسوخة من بحوث منشورة في الباحث العلمي (كوكل سكولر) مع التعديل عليها، وأخرى تحتوي على نصوص منسوخة ملصقة من مقالات منشورة في شبكة الويب.

والرسالة ركّزت على تصميم أداة تتوزع هيكلية عملها إلى أربع مراحل وهي: (الاستدعاء، والتحليل، ومقاييس البيانات، وواجهة المستخدم) ويتم استخدامه لتقييم أداء عدد من البرامج التي تختلف في تقديم خدمة الفحص والكشف عن الانتحال بشكل مجاني أو بطريقة الاشتراك المدفوع.

في حساب التقييم النهائي باستخدام الأداة المصممة بين خمسة برامج لكشف الانتحال حصل برنامج Turnitin على أعلى نسبة قيمة في التقييم (80%) وذلك بعد جمع قيم سبعة مؤشرات مضروبة على حسب معادلات رياضية بقيم أوزان الاستخدام لها وحصل برنامج Small Seo Tools على ثاني اعلى نسبة (65.3%) بحسب تقييم الأداة، والبرنامج iThenticate حصل على نسبة (60%)، والبرنامج DupliChecker جاء بنسبة تقييم (49.8%) وحصل برنامج Grammarly على أقل نسبة في التقييم (44.5%).

وأخيراً، ان تقييم أداء برامج كشف الانتحال له أهمية كبرى، لذا فإن تطوير أداة لتقييم هذا الأداء مع الأخذ بنظر الاعتبار تطور طرائق ممارسة الانتحال والثغرات الموجودة في البرامج المستخدمة لتكون الأداة بجودة عالية في كشف أغلب حالات الانتحال.

Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and
Mathematics
Department of Computer Science

Evalute the Performance of Text Plagiarism Systems Using Software Engineering Measurements

A Thesis Submitted to the Council of the College of Computer Science and Mathematics University of Mosul as a Partial Fulfillment of Requirements for the Degree of Master of Science in Computer Science

By

Nawzt Hasan Ali Hasan

Supervised by

Lectural. Dr. Shayma Mustafa Thanoon

2024 A.D. **1446 A.H.**

Abstract

Plagiarism detection systems and programs are among the most important software applications that serve researchers and students in education and maintain academic integrity.

Because of this, these programs are constantly developed and maintained, and their weak points are addressed to obtain the highest possible benefit from their performance and to prevent the methods used by those who practice plagiarism to question and manipulate the integrity of the research and studies that are published.

In addition, we need to study and evaluate these programs to learn more about the quality of performance and the extent of the possibility of raising the level of the processors they provide.

Many previous works have been studied in evaluating the performance of plagiarism detection programs, the metrics and indicators that go into assessing the performance of these programs, and analyzing software engineering tools to identify the most important differences between the programs. This study included collecting research and work in the field of evaluation from journals and electronic libraries published from the year 2013 until the year 2023.

The research focused on designing a tool whose work structure is divided into four stages: (summoning, analysis, data metrics, and user interface). It is used to evaluate the performance of several programs that differ in providing free plagiarism inspection and detection services or through a paid subscription method.

In calculating the final evaluation using the tool designed among five plagiarism detection programs, the Turnitin program obtained the highest percentage of value in the evaluation (80%) after summing the values of

seven indicators multiplied according to mathematical equations with the values of their usage weights, and the Small Seo Tools program obtained the second highest percentage (65.3%) according to the tool evaluation.

The iThenticate program received a rating of (60%), the DupliChecker program received a rating of (49.8%), and the Grammarly program received the lowest rating.(%44.5)

Finally, evaluating the performance of plagiarism detection programs is of great importance, so developing a tool to evaluate this performance, taking into account the development of methods of practicing plagiarism and the gaps in the programs used so that the tool is of high quality in detecting most cases of plagiarism.