

PERANCANGAN SISTEM INFORMASI IDENTIFIKASI KEMIRIPAN TEKS BERBASIS WEB MENGGUNAKAN ALGORITMA COSINE SIMILARITY

Oleh :

MADE SUSENA SANTIKA

16312018

ABSTRAK

Penggunaan kata-kata atau ide orang lain tanpa memberikan penghargaan kepada penulis aslinya seolah-olah karya atau ide tersebut adalah miliknya sendiri disebut plagiarisme. Selain dilema moral, plagiarisme akan membuat siswa menjadi kurang kreatif dalam memasukkan pemikiran dan pendapatnya ke dalam informasi yang ingin disajikan. Hal ini kemungkinan akan terus dilakukan karena belum adanya sistem yang digunakan untuk mengidentifikasi tingkat kemiripan teks pada instansi terkait. Dengan menilai jurnal yang dapat mengidentifikasi kesamaan beberapa dokumen dengan membandingkan isinya, maka permasalahan plagiarisme pada program studi informatika dapat diatasi. Proses ini menghasilkan nilai atau bobot kemiripan jurnal yang telah dibandingkan. Penelitian ini bertujuan untuk mengembangkan sistem yang dapat mengidentifikasi kemiripan teks secara otomatis dengan memanfaatkan algoritma *cosine similarity*. Untuk mempercepat pengembangan sistem/aplikasi, *Rapid Application Development* (RAD) menggabungkan sejumlah pendekatan terorganisir dengan metodologi pengembangan aplikasi prototipe dan kolaboratif. Paradigma *Rapid Application Development* (RAD) merupakan salah satu inovasi SDLC; dengan menetapkan prioritas waktu, ini juga dapat digunakan untuk pengembangan sistem. Paradigma *Rapid Application Development* (RAD) yang mempunyai fase perancangan dan pengembangan sistem yang relatif singkat (60-90 hari), diterapkan pada sistem yang direncanakan dan diproduksi. Hasil dari penelitian ini adalah pembahasan pada sistem informasi identifikasi kemiripan teks berbasis web menggunakan algoritma *cosine similarity*, disimpulkan bahwa pada penelitian yang dilakukan dapat menghasilkan sebuah sistem informasi yang dapat digunakan untuk mengidentifikasi kemiripan teks pada sebuah artikel dengan menggunakan algoritma *cosine similarity* ini dapat meningkatkan tingkat keakuratan persentase kemiripan teks tersebut dengan artikel pembandingnya.

Kata Kunci : *Algoritma Cosine Similarity, Plagiarisme, Rapid Application Development.*

**DESIGNING A WEB-BASED TEXT SIMILARITY IDENTIFICATION
INFORMATION SYSTEM USING THE COSINE SIMILARITY
ALGORITHM**

Oleh :

MADE SUSENA SANTIKA

16312018

ABSTRACT

Using other people's words or ideas without giving credit to the original author as if the work or idea were his or her own is called plagiarism. Apart from the moral dilemma, plagiarism will make students less creative in inserting their thoughts and opinions into the information they want to present. This will likely continue to be done because there is no system used to identify the level of similarity of texts at the relevant agencies. By assessing journals that can identify similarities between several documents by comparing their contents, the problem of plagiarism in informatics study programs can be overcome. This process produces a similarity value or weight for the journals that have been compared. This research aims to develop a system that can identify text similarities automatically by utilizing the cosine similarity algorithm. To accelerate system/application development, The Rapid Application Development (RAD) paradigm is one of the SDLC innovations; by setting time priorities, this can also be used for system development. The Rapid Application Development (RAD) paradigm, which has a relatively short system design and development phase (60-90 days), is applied to systems that are planned and produced. The result of this research is a discussion of a web-based text similarity identification information system using the cosine similarity algorithm. It is concluded that the research carried out can produce an information system that can be used to identify the similarity of text in an article using the cosine similarity algorithm. This can increase the level of similarity. accurate percentage similarity of the text to the comparison article.

Keywords : Cosine Similarity Algorithm, Plagiarism, Rapid Application Development