

Pengidentifikasian Kesamaan Redaksi Al-Quran Menggunakan Pendekatan Deteksi Plagiat

Galih Yudhasena Trenggala¹, Dr. Moch.Arif Bijaksana, Ir., M. Tech²

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹galihyudha@student.telkomuniversity.ac.id, ²arifbijaksana@telkomuniversity.ac.id

Abstrak

Saat membaca Al-Quran, umat muslim beberapa kali menemukan adanya ayat pada Al-Quran yang berulang atau sama persis saat melafalkannya. Untuk memastikan apakah ayat atau segmen pada Al-Quran mengandung unsur kesamaan atau tidak, dibutuhkan metode yang bisa mengidentifikasi kesamaan redaksi pada Al-Quran. Dalam tugas akhir ini, penulis menggunakan sistem deteksi plagiat atau disebut juga *Text Alignment* untuk mengidentifikasi kesamaan redaksi pada Al-Quran. Penulis menggunakan sistem dari penelitian deteksi plagiat [1], pasangan halaman Al-Quran sebagai masukan pada sistem, dan penggunaan algoritma *Rabin-Karp* untuk pencocokan *string*. Penggunaan pendekatan deteksi plagiat, telah berhasil digunakan dengan mendapatkan performansi (*F1 score*) sebesar 100%.

Kata kunci : al-quran, kesamaan redaksi, deteksi plagiat, *seeding*, *extension*, *filtering*.

Abstract

When reading the Quran, Muslims several times found that there were verses in the Quran that were repeated or the same as they were reciting. To ascertain whether a verse or segment in the Quran contains elements of similarity or not, a method is needed that can identify the editorial similarities in the Quran. In this final project, the author uses a plagiarism detection system or also known as *Text Alignment* to identify editorial similarities in the Quran. The author uses a system from plagiarism detection research, pairs of pages of the Quran as input to the system, and uses the *Rabin-Karp* algorithm for string matching. The use of plagiarism detection approach has been successfully used by obtaining a system performance (*F1 score*) of 100%.

Keywords: al-quran, redaction similiarity, plagiarism, seeding, extension, filtering.

1. Pendahuluan

Latar Belakang

Al-Qur'an merupakan sebuah kitab suci yang diturunkan oleh Allah SWT. melalui malaikat Jibril kepada Nabi Muhammad SAW sebagai petunjuk. Tidak hanya petunjuk untuk umat beragama Islam, tetapi juga untuk seluruh umat manusia yang hidup di bumi. Al-Quran yang terdiri dari 30 juz, 114 surat dan 6236 ayat ini adalah sebuah kitab yang berisi kebaikan dan kemaslahatan untuk umat manusia, baik di dunia maupun akhirat.

Umat muslim beberapa kali menemukan adanya ayat pada Al-Quran yang berulang, bahkan mirip saat melafalkannya. Bagian atau segmen yang berulang ini bisa berada pada halaman, surah dan juz yang berbeda-beda. Untuk memastikan apakah segmen pada Al-Quran mengandung unsur kesamaan atau tidak. Dibutuhkannya metode yang bisa mengidentifikasi kesamaan redaksi pada Al-Quran.

Kesamaan redaksi adalah kesamaan gaya penyusunan kata pada sebuah teks dari dua atau lebih dokumen yang dibandingkan. Untuk mengidentifikasi tingkat kesamaan redaksi pada Al-Quran, ada suatu metode yang bisa digunakan, yaitu metode deteksi plagiat atau disebut juga *Text Alignment*. Metode deteksi plagiat adalah metode yang dipakai untuk mencari kesamaan kata atau kalimat pada dua dokumen atau lebih.

Dari hasil pengidentifikasian ini diharapkan dapat membantu para pengkaji bahasa Al-Quran untuk lebih tepat dalam menemukan kesamaan redaksi pada segmen Al-Quran tanpa adanya kendala kosakata.

Pada penelitian sebelumnya [2], metode deteksi plagiat atau *Text Alignment* telah diuji coba untuk mencari kesamaan pada terjemahan Al-Quran berbahasa Inggris. Maka pada penelitian kali ini, metode deteksi plagiat akan diuji coba untuk mencari pasangan-pasangan halaman Al-Quran yang memiliki kesamaan redaksi dan melihat hasil performansinya.

Batasan Masalah

Adapun batasan dari permasalahan yang ada pada tugas akhir ini sebagai berikut:

1. Mendeteksi segmen Al-Quran yang mempunyai kesamaan redaksi.
2. Dataset yang menjadi masukan merupakan pasangan halaman Al-Quran yang ayat-ayatnya didapat dari situs tanzil.net [3].

Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan di atas, maka permasalahan yang dapat dirumuskan adalah sebagai berikut:

1. Bagaimana mengimplementasikan metode deteksi plagiat untuk mencari segmen yang mempunyai kesamaan redaksi pada Al-Quran?
2. Apakah metode deteksi plagiat dapat menentukan segmen Al-Quran yang memiliki kesamaan redaksi dengan baik?

Tujuan

Berikut adalah tujuan yang ingin dicapai pada tugas akhir ini:

1. Mengimplementasikan metode deteksi plagiat untuk mencari segmen yang mempunyai kesamaan redaksi pada suatu halaman Al-Quran.
2. Melihat hasil performansi sistem dari penggunaan metode deteksi plagiat.

Organisasi Tulisan

Mulai dari bagian 2 akan dijelaskan mengenai studi literatur yang sudah dilakukan. Bagian 3 mengangkat mengenai sistem yang dibangun dalam tugas akhir ini. Pada bagian 4 sistem yang sudah dibangun, akan diuji coba dan hasil dari uji coba tersebut akan dianalisa dan dievaluasi untuk mendapatkan jawaban yang berkaitan dengan masalah yang diangkat dari tugas akhir ini. Bagian 5 berisi kesimpulan hasil evaluasi secara ringkas dan memberikan saran yang dapat dilakukan ke depannya demi memajukan penelitian berikutnya.

2. Studi Terkait

2.1 Plagiarisme

Pada Kamus Besar Bahasa Indonesia, Plagiarisme diartikan sebagai penjiplakan yang melanggar hak cipta. Jadi, deteksi plagiarisme adalah mendeteksi dokumen, tulisan atau karya yang melakukan penjiplakan terhadap dokumen, tulisan atau karya lainnya. Suatu dokumen dikatakan plagiat apabila adanya kesamaan secara sebagian atau menyeluruh dengan dokumen yang lainnya. Plagiarisme sendiri terkategori menjadi lima jenis, yaitu[2]:

1. *no-plagiarism* adalah dokumen yang tidak terdeteksi plagiat sama sekali.
2. *no-obfuscation* adalah dokumen yang terdeteksi plagiat sangat persis tanpa adanya perubahan, penambahan atau pengurangan pada teks.
3. *random-obfuscation* adalah dokumen yang terdeteksi plagiat mirip dengan adanya perubahan, penambahan atau pengurangan pada teks. Jenis plagiarisme *no-obfuscation* tidak digunakan pada tugas akhir ini.
4. *translation-obfuscation* adalah dokumen yang terdeteksi plagiat dengan cara mengubah bahasa ke dalam bahasa lain. Jenis plagiarisme *translation-obfuscation* tidak digunakan pada tugas akhir ini.
5. *summary-obfuscation* yaitu dokumen yang terdeteksi plagiat hampir mirip dengan cara menyimpulkan isi dari dokumen kesatu. Jenis plagiarisme *summary-obfuscation* tidak digunakan pada tugas akhir ini.