

Rancang Bangun Penilaian Keaktifan Menggunakan Teknik *Text Similarity* Pada Sistem Tutorial Matematika Diskret

I Gede Santi Astawa

Program Studi Teknik Informatika, Jurusan Ilmu Komputer
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
e-mail: santi.astawa@cs.unud.ac.id

Luh Putu Ida Harini

Program Studi Matematika, Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
e-mail: ballidah@unud.ac.id

Putri Cahyaning

Program Studi Teknik Informatika, Jurusan Ilmu Komputer
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana

Abstract: Facility of student activeness in adaptif tutorial system expected be one of the supporting facilities. Activity of students in learning process is process that construct their own knowledge. They are actively building an understanding of the issues or anything they encounter in the learning process. Because it is constructed from the inside, causing the activity is difficult to assess without having to do face to face, especially the tendency of students to be silent and conceal their ignorance will cause more difficult to assess the students' activity. In this study, the authors designed an extra feature on discrete mathematics tutorial system that has been built, in the form of an online discussion forum to help educators in scoring liveliness (seen from match students' comments and answer key). In this feature, the method used is the method of measuring similarity (similarity text) Rabin Karp to determine the similarity of the comments. Result of the research shows that the feature has been able to increase student interest in providing comments, which given an average of 2.2 comment in every single student. But rabin Karp algorithm used is still deemed less good if a comment is composed of words with errors in typing, necessitating the addition of other algorithms for increase system performance.

Keywords: *Discussion Forum, text simmilarity, student activity*

Abstrak: Fasilitas penilaian keaktifan mahasiswa dalam sistem tutorial adaptif menjadi salah satu fasilitas penunjang yang diharapkan. Keaktifan siswa dalam kegiatan belajar tidak lain adalah untuk mengkonstruksi pengetahuan mereka sendiri. Mereka aktif membangun pemahaman atas persoalan atau segala sesuatu yang mereka hadapi dalam proses pembelajaran. Karena sifatnya yang terkonstruksi dari dalam diri, menyebabkan keaktifan ini sulit dinilai tanpa harus melakukan tatap muka, apalagi kecenderungan siswa untuk diam dan menyembunyikan ketidakpahaman mereka akan materi menyebabkan semakin sulit menilai keaktifan siswa. Pada penelitian ini penulis merancang sebuah fitur tambahan pada sistem tutorial matematika diskret yang telah dibangun, berupa forum diskusi online untuk membantu pengajar dalam pemberian nilai keaktifan (dilihat dari kecocokan komentar siswa dan kunci jawaban). Dalam fitur ini, metode yang digunakan adalah metode pengukuran kemiripan (*text similarity*) Rabin Karp dengan tujuan mencegah adanya jawaban yang tidak diinginkan guru dengan membandingkan mana jawaban yang sesuai dan mana yang tidak sesuai dan untuk mengetahui kesamaan komentar siswa lainnya. Hasil penelitian menunjukkan bahwa forum diskusi ini telah mampu meningkatkan minat mahasiswa dalam memberikan komentar, dimana dari tiga materi yang diberikan rata-rata setiap mahasiswa memberikan 2,2 komentar. Namun dalam hal penilaian kecocokan komentar, algoritma rabin karp yang digunakan masih dirasa kurang baik apabila sebuah komentar terdiri dari kata-kata yang mengalami kesalahan pengetikan, sehingga diperlukan penambahan algoritma lain untuk memperbaiki kinerja sistem.

Kata kunci: Forum Diskusi, *text similarity*, keaktifan

1. Pendahuluan

Kemajuan teknologi saat ini banyak mendukung berbagai aspek kebutuhan salah satunya dalam memenuhi kebutuhan belajar yang baik. Penggunaan media tutorial berbasis komputer memungkinkan seorang mahasiswa untuk belajar secara mandiri, dengan mengikuti panduan atau arahan-arahan yang terdapat dalam sistem tutorial tersebut. Dalam dunia pendidikan dan pembelajaran, peranan media pembelajaran juga tidak bisa diabaikan. Hasil penelitian Wulan, dkk. [4] menyimpulkan bahwa minat, penguasaan konsep, dan prestasi belajar matematika siswa yang diajar menggunakan media pembelajaran menjadi lebih baik daripada siswa yang diajar menggunakan LKS.

Pada penelitian sebelumnya, telah dilakukan rancang bangun sistem tutorial matematika diskret. Namun dalam sistem tersebut masih memiliki kendala dalam melakukan penilaian keaktifan dari seorang siswa. Keaktifan siswa dalam kegiatan belajar tidak lain adalah untuk mengkonstruksi pengetahuan mereka sendiri. Mereka

aktif membangun pemahaman atas persoalan atau segala sesuatu yang mereka hadapi dalam proses pembelajaran. Karena sifatnya yang terkonstruksi dari dalam diri, menyebabkan keaktifan ini sulit dinilai tanpa harus melakukan tatap muka, apalagi kecenderungan siswa untuk diam dan menyembunyikan ketidakpahaman mereka akan materi menyebabkan semakin sulit menilai keaktifan siswa.

Forum diskusi online merupakan sarana bagi siswa untuk mengungkapkan pemikirannya mengenai sebuah topik diskusi. Melalui forum ini seorang guru mungkin saja dapat menakar pemahaman siswa terhadap materi yang didiskusikan. Dengan forum, siswa akan merasa berdiskusi dengan sesama siswa sehingga akan menjadi lebih berani mengungkapkan apa yang dipikirkannya. Penilaian dapat dilakukan dengan melihat dari kecocokan komentar siswa dengan kunci jawaban yang dimiliki guru untuk satu kali materi yang diberikan. Untuk menentukan kecocokan antara komentar siswa dengan jawaban dari guru, digunakan algoritma Rabin Karp. menurut penelitian dari Atmopawiro tahun 2006, algoritma Rabin Karp ini sangat efektif untuk mencocokkan string yang terdiri dari banyak kata.

Sistem yang dibuat dalam penelitian ini bertujuan untuk memberi penilaian otomatis terhadap komentar siswa berdasarkan kesamaan dengan kunci jawaban yang dimiliki guru. Dalam sistem ini, metode yang digunakan adalah metode pengukuran kemiripan (*text Similarity*) dengan tujuan mencegah adanya jawaban yang tidak diinginkan guru dengan membandingkan mana jawaban yang sesuai dan mana yang tidak sesuai.

2. Tinjauan Pustaka

2.1 Algoritma Rabin Karp

Algoritma Rabin-Karp adalah suatu algoritma pencarian *string* yang diciptakan Michael O. Rabin dan Richard M. Karp pada tahun 1987 yang menggunakan fungsi *hashing* untuk menemukan pola berupa *substring* di dalam sebuah teks. Algoritma ini jarang digunakan untuk melakukan pencarian kata tunggal, namun sangat efektif apabila digunakan untuk pencarian lebih dari satu kata (Atmopawiro [2]).

Dalam algoritma Rabin-Karp, ada beberapa tahap yang harus dilalui dalam implementasi algoritma tersebut.

1) Preprocessing

Tahap ini adalah tahap dimana dilakukan pemrosesan awal untuk mempersiapkan teks agar dapat digunakan dalam proses selanjutnya. Teks akan diubah ke dalam bentuk *k-grams*, dimana dalam hal ini teks akan dipotong-potong dengan panjang *K* dari sebuah teks yang secara kontinu dibaca dari awal dokumen hingga akhir dari dokumen.

2) Rolling Hash

Fungsi yang digunakan untuk menghasilkan nilai hash dari rangkaian *gram* dalam Algoritma Rabin-Karp adalah dengan menggunakan fungsi *rolling hash*.

Fungsi hash $H_{(c_1 \dots c_k)}$ didefinisikan sebagai berikut,

$$c_1 * b^{(k-1)} + c_2 * b^{(k-2)} + \dots + c_{(k-1)} * b^k + c_k \quad (2.1)$$

Keterangan :

c : nilai ASCII karakter

b : basis (bilangan prima)

k : banyak karakter

Metode *hashing* digunakan untuk mempercepat pencarian atau pencocokan suatu *string*. Apabila tidak di-*hash*, pencarian akan dilakukan karakter per karakter. Namun pencarian akan menjadi lebih mangkus setelah di-*hash* karena hanya akan membandingkan empat digit angka untuk mengetahui kesamaan suatu *substring*.

3) Pencocokan

Proses pencocokan dalam Algoritma Rabin-Karp dilakukan dengan menggunakan sebuah teorema yaitu:

"Sebuah *stringA* identik dengan *stringB*, jika (syarat perlu) *stringA* memiliki hash key yang sama dengan hash key yang dimiliki oleh *stringB*"

4) Pengukuran Nilai Kemiripan

Mengukur *similarity* (kemiripan) dan jarak antara dua entitas informasi adalah syarat inti pada semua kasus penemuan informasi, seperti pada *Information Retrieval* dan *Data Mining* yang kemudian dikembangkan dalam bentuk perangkat lunak, salah satunya adalah sistem deteksi kesamaan (Eko [3]). Penggunaan ukuran *similarity* yang tepat tidak hanya meningkatkan kualitas pilihan informasi tetapi juga membantu mengurangi waktu dan biaya proses sehingga memperangkat lunakkan *Dice's Similarity Coefficient* dalam penghitungan nilai *similarity* yang menggunakan pendekatan *k-gram*.

$$S = \frac{K \times C}{(A + B)}$$

Dimana S adalah nilai *similarity*, A dan B adalah jumlah dari kumpulan *k-grams* dalam teks 1 dan teks 2. C adalah jumlah dari *k-grams* yang sama dari teks yang dibandingkan.

2.2 Text Preprocessing

Melihat dari kemungkinan berbagai macam bentuk struktur data teks yang akan diproses, maka sebelum dilakukannya penghitungan kesamaan teks perlu dilakukan beberapa tahap awal yang bertujuan untuk agar teks dapat diubah menjadi

lebih terstruktur dan menghilangkan *noise* pada dokumen. Proses *preprocessing* tersebut meliputi *case folding*, *tokenizing*, *filtering*, *stemming*, dan *pelabelan kata* (Anna [1]).

a) Case Folding

Case folding merupakan proses pertama dari rangkaian *preprocessing* dokumen. Dalam proses ini akan dilakukan perubahan pada kata-kata dalam dokumen menjadi huruf kecil (a sampai z). Berikut ini merupakan contoh dari proses *case folding* :

[1] *Input* : Rabin Karp adalah algoritma pencarian kata yang mencari sebuah pola dalam sebuah teks.

[2] *Output* : rabin karp adalah algoritma pencarian kata yang mencari sebuah pola dalam sebuah teks.

b) Tokenizing

Tokenizing merupakan tahapan dimana dilakukannya pemotongan terhadap *stringinput* berdasarkan atas *delimiter* yang telah ditentukan. Karakter selain huruf akan dianggap sebagai *delimiter* dan akan dihilangkan atau dihapus untuk proses mendapat kata-kata penyusun teks. Dari proses ini akan dihasilkan kata-kata penyusun *string/teks* atau yang sering disebut *token/term*. Berikut ini merupakan contoh dari proses *tokenizing*:

[1] *Input* : rabin karp adalah algoritma pencarian kata yang mencari sebuah pola dalam sebuah teks.

[2] *Output* : {rabin} {karp} {adalah} {algoritma} {pencarian} {kata} {yang} {mencari} {sebuah} {pola} {dalam} {sebuah} {teks}

c) Filtering

Filtering merupakan tahap pengambilan kata-kata penting dari hasil *tokenizingstring*. *Filtering* dilakukan dengan membuang kata-kata yang telah terdaftar ke dalam *stopword/stoplist*. *Stopword* adalah kata-kata yang sering muncul dalam teks dalam jumlah besar dan dianggap tidak memiliki makna penting. Berikut ini merupakan contoh dari proses *filtering* :

[1] *Input* : {rabin} {karp} {adalah} {algoritma} {pencarian} {kata} {yang} {mencari} {sebuah} {pola} {dalam} {sebuah} {teks}

[2] *Output* : {rabin} {karp} {algoritma} {pencarian} {kata} {mencari} {pola} {teks}

d) Stemming

Stemming merupakan proses yang dilakukan untuk mendapatkan kata dasar dari suatu kata. *Stemming* Nazief – Adriani merupakan suatu algoritma *stemming* yang dibuat oleh Bobby Nazief dan Mirna Adriani.

Pencarian kata dasar dilakukan dengan menghilangkan semua imbuhan dari kata, baik itu awalan (*prefiks*), sisipan (*infiks*), maupun akhiran (*suffiks*).

Berikut ini merupakan contoh dari proses *stemming* :

[1] Input : {rabin} {karp} {algoritma} {pencarian} {kata} {mencari} {pola} {teks}

[2] Output : {rabin} {karp} {algoritma} {cari} {kata} {cari} {pola} {teks}

3. Implementasi Sistem

Pemberian nilai keaktifan ini dapat diakses oleh 3 jenis pengguna yaitu siswa, guru dan admin. Sebelum dapat mengakses sistem dan menjalankan fungsinya masing-masing, pengguna terlebih dahulu harus melakukan login.

Untuk melakukan login, pengguna harus memasukkan *username* dan *password*.



Gambar 1. Antarmuka Forum Diskusi

Gambar 1 di atas menampilkan halaman Forum Diskusi Siswa. Pada antarmuka ini terdapat pilihan materi sesuai pelajaran yang diambil di kelas.



Gambar 2. Antarmuka Lihat Materi Diskusi

Gambar 2 di atas merupakan antar muka menampilkan deskripsi atau soal yang diberikan guru, setelah user siswa memilih materi yang bersangkutan.



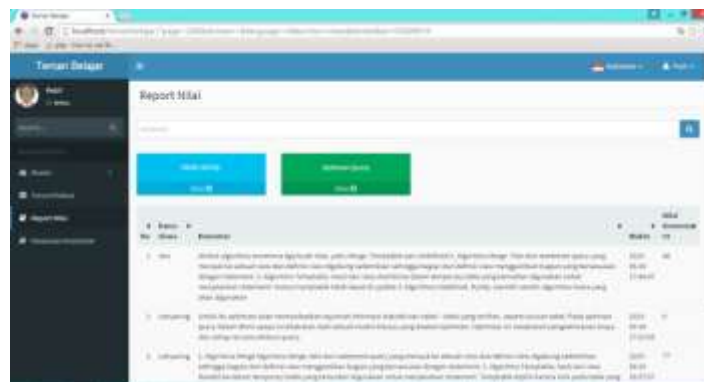
Gambar 3. Antarmuka Submit Komentar

Gambar 3 antarmuka di atas menampilkan beberapa komentar siswa dari diskusi yang diberikan Guru. Komentar yang di submit kemudian akan disimpan di *database* yang kemudian akan dihitung *similarity*nya dengan kunci jawaban guru.



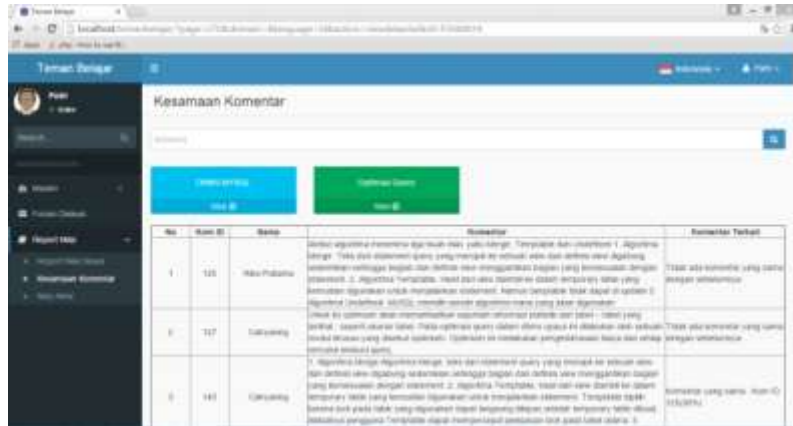
Gambar 4. Antarmuka Tambah Materi Diskusi

Gambar 4 antarmuka di atas adalah form user guru untuk menambah bahan materi. tabel yang terdapat pada forum diskusi dengan tambahan menu edit dan delet data materi.



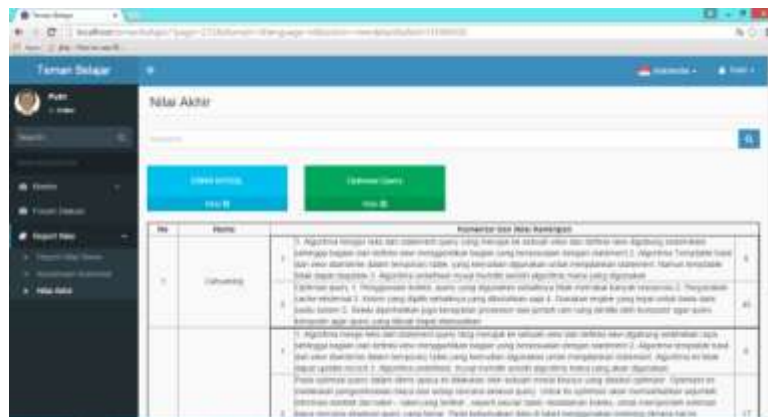
Gambar 5. Antarmuka Hasil Report Similarity Komentar dengan Jawaban Guru

Gambar 5 di atas menunjukkan tampilan hasil dari perhitungan *similarity* oleh sistem. Tabel akan menampilkan nama siswa yang memberikan komentar, komentar atau jawaban yang diberikan, waktu dan tanggal pemberian komentar di Forum Diskusi Siswa dan Nilai. Pada kolom nilai komentar, nilai yang dihasilkan berupa hasil dari perhitungan *similarity*.



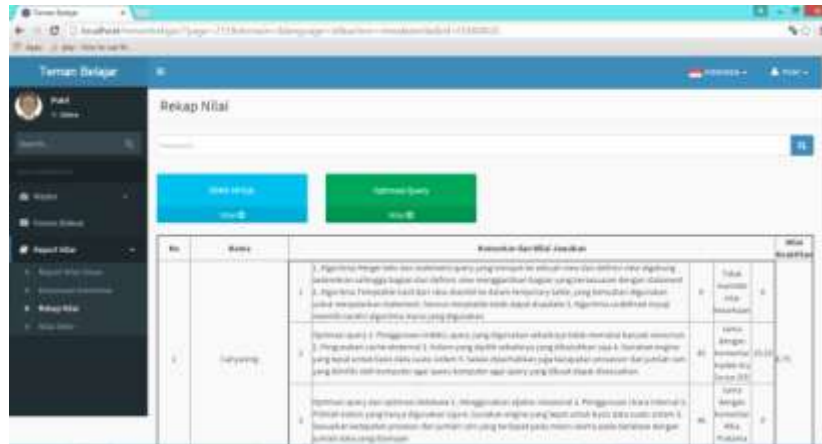
Gambar 6. Antarmuka untuk menghitung Kesamaan Komentar

Dari komentar yang diberikan, akan dihitung lagi kemiripan dari komentar tersebut. tujuannya untuk mengetahui adanya kesamaan jawaban yang dimiliki setiap siswa dalam pemberian komentar disetiap materi diskusi atau soal yang diberikan.



Gambar 7. Antarmuka untuk Menampilkan Nilai Tertinggi

Dari gambar 7 di atas menjelaskan, guru bisa memilih menu untuk menampilkan daftar nilai tertinggi dari siswa yang memberikan komentar dengan nilai tertinggi, agar bisa dijadikan laporan akhir dari penilaian keaktifan.

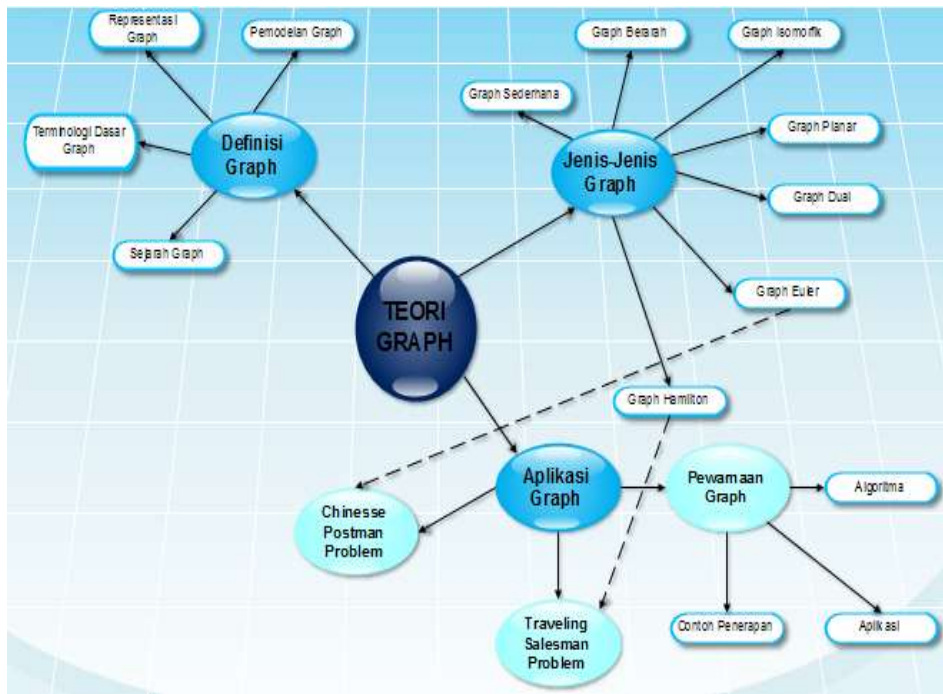


Gambar 8. Antarmuka Menentukan Nilai Keaktifan Siswa

Antarmuka gambar 8 di atas menjelaskan bagaimana sistem menentukan perhitungan nilai keaktifan. Nilai keaktifan didapat dari hasil similarity jawaban dengan kunci jawaban dan juga dihitung dari nilai similarity sesama jawaban siswa.

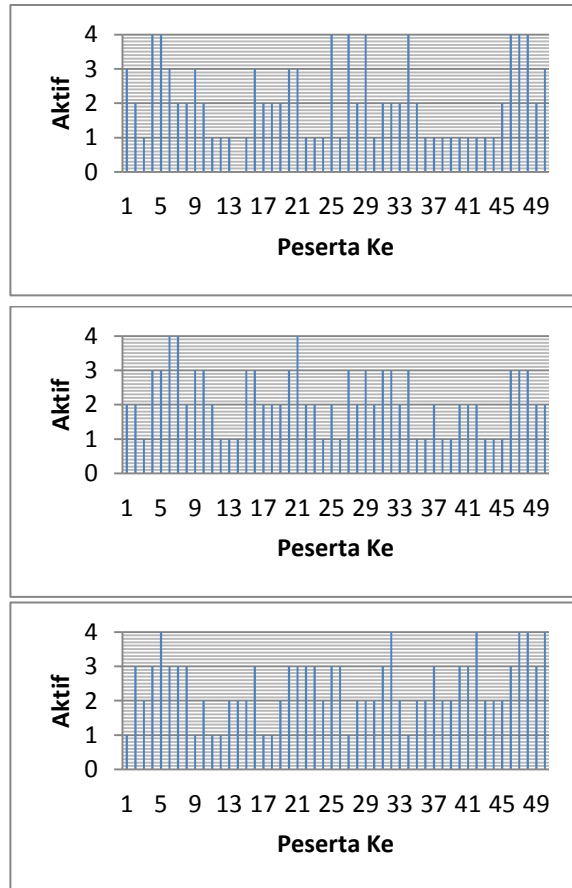
4. Pembahasan

Selanjutnya digunakan materi graf dalam matematika diskret sebagai studi kasus dalam pengujian sistem penilai keaktifan yang telah dibangun. Gambar 9. menunjukkan keterkaitan antar materi dalam graft.



Gambar 9. Keterkaitan Materi dalam Graft

Pengujian dilakukan dengan penerapam langsung kepada 50 mahasiswa program studi teknik informatika yang sedang mengambil matakuliah matematika diskret. Materi dalam pengujian ini ditetapkan materi chininese postman problem, materi travelling salesman problem, dan materi pewarnaan graft. Dari tiga materi tersebut diperoleh keikutsertaan mahasiswa sebagai berikut:



Gambar 10. Sebaran Keaktifan Peserta dalam Memberi Komentar Pada Tiga Materi Diskusi

Dari gambar 10 di atas, rata-rata keikutsertaan mahasiswa dalam media forum diskusi ini adalah 2, 2 komentar per mahasiswa, sehingga dapat dikatakan bahwa partisipasi peserta mahasiswa cukup baik. Disamping itu terlihat pula pada gambar 10 bahwa hampir seluruh peserta mahasiswa memberikan komentar pada setiap materi yang diberikan.

Pengujian selanjutnya dilakukan untuk melihat kinerja sistem dalam menilai kecocokan komentar dengan jawaban yang diharapkan pengajar. Skenario pengujian adalah dengan memberikan dua alternatif jawaban pada masing-masing materi diskusi dan melihat tingkat similitasnya berdasarkan perhitungan algoritma rabin

karp. Kemudian dilakukan analisa dan perbandingan dari 5 komentar dengan nilai similaritas terendah. Berikut analisa dari komentar dengan nilai similaritas terkecil

Komentar “menurut saya pewarnaan graf adalah algoritma untuk mencari bentuk dari model bilangan kromatik sebuah graf”

Alternatif Jawaban

“pewarnaan graft merupakan algoritma untuk menentukan jumlah warna minimum yang diperlukan apabila setiap node yang terhubung dalam graft tidak boleh memiliki warna yang sama”

“pewarnaan graft adalah cara mencari nilai bilangan kromatik sebuah graft, yaitu warna minimum yang diperlukan untuk menandai node pada graft dengan syarat node yang terhubung tidak boleh memiliki warna yang sama”

Dari komentar siswa dan alternatif jawaban di atas, terlihat bahwa komentar siswa hanya mengandung 5 kata yang terkandung dalam alternatif jawaban, yaitu “algoritma”, “mencari”, “bilangan”, ”kromatik”, ”sebuah”. Kata-kata kunci yang lain ternyata mengandung kesalahan penulisan seperti **“grafft”** Hal ini menyebabkan similaritasnya menjadi sangat kecil, walaupun apabila dibaca secara manual komentar tersebut cukup baik.

5. Kesimpulan dan Saran

Kesimpulan yang dapat diambil dari penelitian ini adalah: (1) Minat peserta mahasiswa dalam berperan serta aktif melalui forum diskusi ini cukup tinggi dan hampir seluruh peserta memberikan komentar; (2) Algoritma rabin karp memberikan penilaian berdasarkan kesamaan kata, sehingga menimbulkan kesalahan penilaian apabila terjadi kesalahan penulisan pada komentar mahasiswa.

Berdasarkan hasil penelitian, untuk pengembangan sistem kedepannya sebaiknya lebih disempurnakan dalam pemeriksaan kembali pada ketepatan kata yang digunakan, seperti penggunaan algoritma *autocorrect*.

Ucapan Terima Kasih

Atas terlaksananya penelitian ini, penulis mengucapkan terima kasih kepada Dirjen Dikti melalui Hibah Penelitian (Hibah Bersaing) Anggaran Tahun 2015, dengan kontrak nomor: 311-37/UN14.2/PNL.01.03.00/2015. Hasil penelitian tidak akan bermanfaat jika tidak dipublikasikan, untuk itu ucapan terima kasih juga disampaikan kepada redaksi jurnal Matematika Universitas Udayana, sehingga artikel ini dapat dipublikasikan.

Daftar Pustaka

- [1] Anna K, Kemal A, S, I wayan S, W. 2012. Jurusan Sistem Informasi Fakultas Ilmu komputer dan Teknologi Informasi Universitas Gunadarma. *Arsitektur Untuk Aplikasi Deteksi Kesamaan Dokumen Bahasa Indonesia*.
- [2] Atmopawiro, A. 2006. Pengkajian dan Analisis Tiga Algoritma Efisien Rabin Karp, Knuth-Morris-Pratt, dan Boyer Moore dalam Pencarian Pola Dalam Suatu Teks.
- [3] Eko G, H, Yus G, V, Anggia F, M. 2013. Departemen Teknik Informatika, Institut Teknologi Bandung Analisis Algoritma Pencarian String (*String Matching*)
- [4] Wulan, E.R., dkk. 2004. *Keefektifan Penggunaan Media Model dalam Pembelajaran Matematika*. Jurnal Teknodika 2 (3) : 44-63.