

**SOAL UJIAN AKHIR SEMESTER GENAP 2020-2011****Program Studi Magister Teknik Informatika****Fakultas Ilmu Komputer
Universitas ESA UNGGUL**

| | | | |
|--------------------------------|-----------------------------------|--|----------|
| Kode/Mata Kuliah | : TOPIK DALAM PEMROGRAMAN | | |
| Dosen | : Dr. BUDI TJAHJONO, S.Kom, M.Kom | | |
| Hari | : SABTU | Waktu | : 12 JAM |
| Tanggal | : 7 AGUSTUS 2021 | Seksi | : KJ101 |
| SifatUjian | : OPEN BOOK | | |
| Kolom Verifikasi Soal | | | |
| Tanggal dan Tanda Tangan Dosen | | Tanggal dan Tanda Tangan Ketua Program Studi | |
| | | | |

KERJAKAN SOAL BERIKUT DENGAN TELITI DAN SESUAI DENGAN PERINTAH!

1. Soal berkaitan dengan tugas besar pembuatan jurnal:
 - a. Apakah masalah yang melatar belakangi anda membuat jurnal tersebut?

JAWAB

Perkembangan teknologi informasi semakin meningkat dan berdampak positif. Salah satu dampak positif dari perkembangan teknologi informasi adalah kemudahan dalam pertukaran informasi. Kemudahan ini sering disalahgunakan oleh seseorang atau beberapa orang dalam menyelesaikan pekerjaan. Penyalahgunaan biasanya terjadi pada Dokumen Ilmiah. Dokumen ilmiah adalah dokumen penalaran dan hasil penelitian dengan menggunakan berbagai objek dan metode. Dokumen yang baik memiliki referensi dan sumber yang eksplisit sebagai referensi, misalnya karya ilmiah atau jurnal yang memiliki referensi dan kutipan eksplisit serta tidak menyalahgunakan kutipan asli dokumen tersebut.

Naskah jurnal ilmiah adalah naskah penelitian ilmiah penulis dan terdiri dari kutipan langsung dan tidak langsung. Kutipan langsung adalah kutipan yang penulisannya sama persis dengan sumbernya, baik dalam bahasa maupun ejaan. Kutipan tidak langsung adalah kutipan yang tidak sama dengan aslinya, dan pengutip hanya mengambil ide pokok dari sumber yang dikutip untuk disajikan kembali dengan kalimat yang disusun oleh kutipan tersebut. Kutipan dianggap penting sebagai referensi yang sah dan dicantumkan dalam penulisan karya ilmiah.

Kesalahan yang merugikan penulis naskah jurnal ilmiah disebut plagiarisme. Berdasarkan penelitian sebelumnya yang menggunakan metode cosine similarity untuk mengecek plagiarisme saja, penelitian ini melengkapi dan menyempurnakannya dengan menimbang setiap kata dalam dokumen ilmiah dan menguji nilai kemiripan teks. Jadi, dokumen yang memiliki nilai kemiripan yang tinggi adalah tanda kutip yang disalahgunakan dan nilai kemiripan yang tinggi. Kemiripan yang rendah adalah dokumen yang telah dikutip dengan sedikit penyalahgunaan atau menuju kebenaran.

Sistem pengecekan kutipan referensi bibliografi berbasis website untuk jurnal ilmiah akan digunakan dalam pengecekan dokumen ilmiah. Kemudian, penelitian ini menggunakan metode pembobotan TF-IDF yang berfungsi untuk menentukan

nilai pembobotan validitas naskah jurnal ilmiah, dan Cosine Similarity Algorithm yang mencari persamaan kata sesuai naskah jurnal ilmiah.

Prinsip dasar metode yang digunakan dalam penelitian ini adalah Text Preprocessing menggunakan Case Folding, Tokenizing, Filtering, Stemming. Selanjutnya nilai bobot kata dalam dokumen ilmiah yaitu TF-IDF dilakukan setelah stemming. Selain itu, tahapan penghapusan stopwords dihitung berdasarkan nilai atau bobot sebuah kata (term) dalam dokumen. Setelah didapatkan nilai bobot kata, selanjutnya dilakukan pengujian tingkat kemiripan menggunakan Cosine similarity, yaitu membandingkan kemiripan antar dokumen. Dalam hal ini yang dibandingkan adalah query dengan dokumen training. Dalam menghitung cosinus similarity, hal pertama yang harus dilakukan adalah melakukan perkalian skalar antara query dan dokumen, menjumlahkannya, mengalikan panjang dokumen dan kuadrat dari panjang query, kemudian menghitung akar kuadrat. Selanjutnya hasil perkalian skalar dibagi dengan hasil perkalian panjang dokumen dan query. Dalam pekerjaan terkait, metode ini banyak digunakan dan menunjukkan kinerja yang baik dalam beberapa kasus. Selain itu, dimungkinkan untuk menambahkan peningkatan apa pun untuk meningkatkan akurasi dan kinerjanya

b. Tunjukkan bagaimana akan menyelesaikan masalah tersebut?

1) Text Processing

Penambangan teks atau dengan nama lain seperti analisis teks cerdas, penambangan data teks, atau penemuan pengetahuan dalam sebuah teks secara sederhana dapat diartikan sebagai proses menemukan pola yang sebelumnya tidak terlihat dalam dokumen atau sumber teks tertentu.

Langkah awal dari text mining adalah preprocessing. Beberapa di antaranya adalah case folding atau pengubahan teks menjadi huruf kecil dan tanpa tanda baca atau karakter khusus, filtering yaitu dengan menghilangkan konjungsi, konjungsi, tokenizing pemisahan dokumen atau kalimat ke dalam setiap kata atau istilah, dan Stemming atau pemisahan imbuhan pada kata menjadi kata dasar.

Langkah selanjutnya adalah tokenizing. Proses ini mengubah kalimat menjadi kata-kata dengan spasi sebagai fitur pemisah

Pemfilteran adalah menghapus kata-kata yang tidak perlu dalam teks input. Jadi, teks yang difilter akan mudah diproses di tahap selanjutnya. Penelitian ini menggunakan 758 kata yang disimpan dalam dataset eksternal. Setelah itu stemming. Langkah ini mengubah kata yang disaring menjadi kata dasar dan menghilangkan imbuhan.

2) TF-IDF and Cosine Similarity

TF diperoleh dari jumlah kata yang muncul dalam sebuah teks, sehingga Anda dapat mengetahui berapa kali kata tersebut muncul. Dan setelah perhitungan TF ini akan dilanjutkan dengan perhitungan DF. Tabel 4 merupakan proses penghitungan nilai bobot untuk menentukan jumlah kata yang muncul pada teks kutipan dan sumber kutipan sehingga dapat dihitung dengan menghitung cosinus similarity

Kemudian menghitung nilai DF (document frequency), yaitu banyaknya kata yang muncul pada kedua dokumen yang diuji

Setelah melakukan perhitungan IDF (inverse document frequency), langkah selanjutnya adalah menganalisa IDF, nilai TF, dan nilai IDF yang sudah didapatkan dari perhitungan diatas, selanjutnya analisa TFIDF merupakan perhitungan untuk mendapatkan nilai bobot dari kutipan teks dan sumber kutipan yang diuji.

Perhitungan cosinus similarity dijelaskan pada tabel 8 dimana dokumen 1 adalah teks kutipan dan dokumen 2 adalah teks sumber kutipan. Teks kutipan dikalikan

dengan teks sumber kutipan, teks kutipan dan sumber kutipan dikuadratkan sehingga diperoleh total untuk mencari nilai kemiripan teks.

Berdasarkan hasil perhitungan algoritma manual di atas (4), nilai kesamaan pada teks adalah 0,4036. Dalam penelitian ini tingkat kemiripan akan dibagi menjadi 3 kategori yaitu tinggi, sedang, dan rendah. Level rendah berada pada kisaran 0-0,3, level sedang berada pada kisaran 0,3-0,6, dan level tinggi berada pada kisaran 0,6-1,0. Hasil (4) adalah 0,4036 sehingga diketahui teks tersebut memiliki tingkat sedang.

- c. Apabila anda telah menyebutkan cara menyelesaikan masalah tersebut, manfaat apa yang anda harapkan dengan anda melakukan penelitian dalam menyelesaikan masalah tersebut?

JAWAB

Kutipan memiliki peran penting dalam artikel ilmiah. Setiap kalimat yang dicantumkan harus diperoleh dari referensi. Ada ukuran yaitu kesamaan. Masalahnya, kemiripan kalimat dalam artikel ilmiah dengan sumber yang dirujuk bisa berbeda makna. Tingkat kemiripan yang tinggi tentunya dapat meningkatkan nilai kemiripan jika dicek menggunakan alat plagiarisme (similarity checker), dimana semakin tinggi nilainya semakin mirip dengan artikel lain sebaliknya dengan sisa kemiripan yang rendah. Namun, tingkat kemiripan yang rendah juga dapat berarti bahwa suatu pernyataan dalam karya ilmiah telah merujuk pada sumber yang sama sekali tidak relevan. Penelitian ini bertujuan untuk menganalisis tingkat kemiripan antara artikel ilmiah dengan artikel atau sumber lain yang dirujuk. Pengecekan teks input dimulai dengan preprocessing teks, pembobotan TF-IDF, dan penentuan tingkat kemiripan menggunakan Cosine Similarity. Berdasarkan hasil pengujian diketahui bahwa hasil dokumen artikel ilmiah memiliki tingkat kemiripan yang tinggi dengan artikel yang dirujuk, dan 30% lainnya memiliki tingkat kemiripan yang rendah, atau dengan kata lain tidak ada hubungan.

- d. Apakah tujuan anda menyelesaikan masalah tersebut?

JAWAB

Kutipan memiliki peran penting dalam artikel ilmiah. Setiap kalimat yang dicantumkan harus diperoleh dari referensi. Ada ukuran yaitu kesamaan. Masalahnya, kemiripan kalimat dalam artikel ilmiah dengan sumber yang dirujuk bisa berbeda makna. Tingkat kemiripan yang tinggi tentunya dapat meningkatkan nilai kemiripan jika dicek menggunakan alat plagiarisme (similarity checker), dimana semakin tinggi nilainya semakin mirip dengan artikel lain sebaliknya dengan sisa kemiripan yang rendah. Namun, tingkat kemiripan yang rendah juga dapat berarti bahwa suatu pernyataan dalam karya ilmiah telah merujuk pada sumber yang sama sekali tidak relevan. Penelitian ini bertujuan untuk menganalisis tingkat kemiripan antara artikel ilmiah dengan artikel atau sumber lain yang dirujuk. Pengecekan teks input dimulai dengan preprocessing teks, pembobotan TF-IDF, dan penentuan tingkat kemiripan menggunakan Cosine Similarity. Berdasarkan hasil pengujian diketahui bahwa hasil dokumen artikel ilmiah memiliki tingkat kemiripan yang tinggi dengan artikel yang dirujuk, dan 30% lainnya memiliki tingkat kemiripan yang rendah, atau dengan kata lain tidak ada hubungan.

- e. Menggunakan metodologi apakah anda menyelesaikan masalah tersebut?

JAWAB

- 1) Text Processing
- 2) TF-IDF and Cosine Similarity

- f. Bagaimana anda melakukan penelitian dengan menggunakan metodologi yang telah ditetapkan untuk menyelesaikan masalah?

JAWAB

- 1) Text Processing

Penambangan teks atau dengan nama lain seperti analisis teks cerdas, penambangan data teks, atau penemuan pengetahuan dalam sebuah teks secara sederhana dapat diartikan sebagai proses menemukan pola yang sebelumnya tidak terlihat dalam dokumen atau sumber teks tertentu.

Langkah awal dari text mining adalah preprocessing. Beberapa di antaranya adalah case folding atau pengubahan teks menjadi huruf kecil dan tanpa tanda baca atau karakter khusus, filtering yaitu dengan menghilangkan konjungsi, konjungsi, tokenizing pemisahan dokumen atau kalimat ke dalam setiap kata atau istilah, dan Stemming atau pemisahan imbuhan pada kata menjadi kata dasar.

Langkah selanjutnya adalah tokenizing. Proses ini mengubah kalimat menjadi kata-kata dengan spasi sebagai fitur pemisah

Pemfilteran adalah menghapus kata-kata yang tidak perlu dalam teks input. Jadi, teks yang difilter akan mudah diproses di tahap selanjutnya. Penelitian ini menggunakan 758 kata yang disimpan dalam dataset eksternal. Setelah itu stemming. Langkah ini mengubah kata yang disaring menjadi kata dasar dan menghilangkan imbuhan.

- 2) TF-IDF and Cosine Similarity

TF diperoleh dari jumlah kata yang muncul dalam sebuah teks, sehingga Anda dapat mengetahui berapa kali kata tersebut muncul. Dan setelah perhitungan TF ini akan dilanjutkan dengan perhitungan DF. Tabel 4 merupakan proses penghitungan nilai bobot untuk menentukan jumlah kata yang muncul pada teks kutipan dan sumber kutipan sehingga dapat dihitung dengan menghitung cosinus similarity

Kemudian menghitung nilai DF (document frequency), yaitu banyaknya kata yang muncul pada kedua dokumen yang diuji

Setelah melakukan perhitungan IDF (inverse document frequency), langkah selanjutnya adalah menganalisa IDF, nilai TF, dan nilai IDF yang sudah didapatkan dari perhitungan diatas, selanjutnya analisa TFIDF merupakan perhitungan untuk mendapatkan nilai bobot dari kutipan teks dan sumber kutipan yang diuji.

Perhitungan cosinus similarity dijelaskan pada tabel 8 dimana dokumen 1 adalah teks kutipan dan dokumen 2 adalah teks sumber kutipan. Teks kutipan dikalikan dengan teks sumber kutipan, teks kutipan dan sumber kutipan dikuadratkan sehingga diperoleh total untuk mencari nilai kemiripan teks.

Berdasarkan hasil perhitungan algoritma manual di atas (4), nilai kesamaan pada teks adalah 0,4036. Dalam penelitian ini tingkat kemiripan akan dibagi menjadi 3 kategori yaitu tinggi, sedang, dan rendah. Level rendah berada pada kisaran 0-0,3, level sedang berada pada kisaran 0,3-0,6, dan level tinggi berada pada kisaran 0,6-1,0. Hasil (4) adalah 0,4036 sehingga diketahui teks tersebut memiliki tingkat sedang.

- g. Jelaskan tahapan anda melakukan penelitian!

JAWAB

1) Text Processing

Penambangan teks atau dengan nama lain seperti analisis teks cerdas, penambangan data teks, atau penemuan pengetahuan dalam sebuah teks secara sederhana dapat diartikan sebagai proses menemukan pola yang sebelumnya tidak terlihat dalam dokumen atau sumber teks tertentu.

Langkah awal dari text mining adalah preprocessing. Beberapa di antaranya adalah case folding atau pengubahan teks menjadi huruf kecil dan tanpa tanda baca atau karakter khusus, filtering yaitu dengan menghilangkan konjungsi, konjungsi, tokenizing pemisahan dokumen atau kalimat ke dalam setiap kata atau istilah, dan Stemming atau pemisahan imbuhan pada kata menjadi kata dasar.

Langkah selanjutnya adalah tokenizing. Proses ini mengubah kalimat menjadi kata-kata dengan spasi sebagai fitur pemisah

Pemfilteran adalah menghapus kata-kata yang tidak perlu dalam teks input. Jadi, teks yang difilter akan mudah diproses di tahap selanjutnya. Penelitian ini menggunakan 758 kata yang disimpan dalam dataset eksternal. Setelah itu stemming. Langkah ini mengubah kata yang disaring menjadi kata dasar dan menghilangkan imbuhan.

2) TF-IDF and Cosine Similarity

TF diperoleh dari jumlah kata yang muncul dalam sebuah teks, sehingga Anda dapat mengetahui berapa kali kata tersebut muncul. Dan setelah perhitungan TF ini akan dilanjutkan dengan perhitungan DF. Tabel 4 merupakan proses penghitungan nilai bobot untuk menentukan jumlah kata yang muncul pada teks kutipan dan sumber kutipan sehingga dapat dihitung dengan menghitung cosinus similarity

Kemudian menghitung nilai DF (document frequency), yaitu banyaknya kata yang muncul pada kedua dokumen yang diuji

Setelah melakukan perhitungan IDF (inverse document frequency), langkah selanjutnya adalah menganalisa IDF, nilai TF, dan nilai IDF yang sudah didapatkan dari perhitungan diatas, selanjutnya analisa TFIDF merupakan perhitungan untuk mendapatkan nilai bobot dari kutipan teks dan sumber kutipan yang diuji.

Perhitungan cosinus similarity dijelaskan pada tabel 8 dimana dokumen 1 adalah teks kutipan dan dokumen 2 adalah teks sumber kutipan. Teks kutipan dikalikan dengan teks sumber kutipan, teks kutipan dan sumber kutipan dikuadratkan sehingga diperoleh total untuk mencari nilai kemiripan teks.

Berdasarkan hasil perhitungan algoritma manual di atas (4), nilai kesamaan pada teks adalah 0,4036. Dalam penelitian ini tingkat kemiripan akan dibagi menjadi 3 kategori yaitu tinggi, sedang, dan rendah. Level rendah berada pada kisaran 0-0,3, level sedang berada pada kisaran 0,3-0,6, dan level tinggi berada pada kisaran 0,6-1,0. Hasil (4) adalah 0,4036 sehingga diketahui teks tersebut memiliki tingkat sedang.

- h. Bila terdapat rumus2 dan Langkah-langkah yang spesifik, tunjukkan langkahnya!

JAWAB

Tidak ada

- i. Apakah hasil penelitian anda sesuai dengan manfaat yang diharapkan?

JAWAB

Iya hasil penelitian sesuai dengan manfaat yang diharapkan.

- j. Yang manakah hasil penelitian yang sesuai, dan manakah yang tidak sesuai?

JAWAB

Skenario pengujian sistem mendefinisikan jumlah data yang digunakan dan proses apa yang dilakukan dalam sistem untuk menentukan kesalahan dalam kutipan jurnal ilmiah ini. Proses yang dilakukan dalam pengujian sistem ini adalah pengumpulan data, text preprocessing, weighting, dan text similarity Cosine Similarity. Data yang digunakan dalam pengujian sistem ini adalah jurnal online yang tersedia di internet dengan mencari referensi daftar pustaka yang tertera pada dokumen yang akan diuji dan menginput teks kutipan. 10 dari 50 sampel uji dapat dilihat pada tabel di bawah ini.

Pada tabel di atas terdapat ambang batas yaitu jika kemiripan teks 0-0,3 berarti tingkat kemiripannya rendah. Jika kemiripan teks adalah 0,3-0,6 maka kemiripannya sedang. Jika teks kesamaan menghasilkan nilai 0,6-1,0, tingkat kesamaan tinggi seperti yang dijelaskan pada bagian metodologi.

Berdasarkan 50 sampel data yang diuji memiliki nilai kemiripan yang berbeda untuk setiap input teks. Hal yang mempengaruhi nilai kemiripan adalah jumlah bobot kata yang diuji. Pada sampel data pertama, jumlah teks 1 memiliki 9 kata dan teks input 2 memiliki 15 kata, dengan jumlah bobot kata yang tidak terlalu jauh dalam jangkauan. Pada pengujian teks input kedua, antara teks 1 dan teks 2 jumlah bobot kata tidak proporsional, sehingga didominasi oleh teks 2. Dan nilai kemiripan diuji berdasarkan bobot kata yang muncul pada kedua input teks. Dokumen dengan tingkat sedang dan tinggi mengenai kesesuaian kutipan dan sumber kutipan adalah 70%, dan tingkat rendah 30%, dokumen tidak setara adalah dokumen yang tidak memiliki kesesuaian antara kutipan dan sumber kutipan. Hasil pengecekan validasi data dokumen jurnal ilmiah adalah 70% sitasi dalam artikel adalah benar dan 30% sitasi salah.

- k. Berikan saranj untuk penelitian selanjutnya.

JAWAB

Memeriksa kutipan jurnal ilmiah dalam menguji data sampel, ditemukan teks yang tidak serupa. Teks kutipan yang tidak memiliki nilai kemiripan, teks yang mendekati kemiripan adalah teks yang memiliki nilai mendekati kemiripan, dan teks serupa adalah teks yang memiliki nilai kemiripan. Pembobotan TF-IDF dalam menguji satu teks kutipan dengan satu teks sumber kutipan yang berasal dari daftar pustaka, maka pembobotan sangat mempengaruhi nilai kemiripan teks. Kesamaan Cosine Similarity dalam menguji satu teks kutipan dengan satu teks sumber kutipan yang berasal dari daftar pustaka. Kemudian nilai kemiripan teks diperoleh nilai ambang batas dan mengetahui teks termasuk dalam beberapa kategori yaitu tidak mirip, hampir mirip, dan mirip dengan akurasi serupa dan Mendekati serupa, yaitu 70% dan tidak 30% mirip karena kesamaan dan mendekati serupa. Ada kesepakatan antara kutipan dan sumbernya dan nilai bobot kata yang mempengaruhinya.

2. Soal berkaitan dengan materi perkuliahan:

- a. Berikan rincian singkat tentang materi perkuliahan yang anda pelajari selama satu semester!

JAWAB

- 1) Berpikir Komputasional
- 2) Bahasa Pemrograman dan Sejarahnya
- 3) Pemrograman Fungsional
- 4) Pemrograman Prolog

- 5) Pendekatan berorientasi objek
- 6) Apa itu pemrograman even driven
- 7) Paradigma pemrograman
- 8) Memahami paradigma pemrograman procedural
- 9) Enam tipe struktur data utama
- 10) Kompleksitas Algoritma
- 11) Algoritma Brute Force dalam pencarian String
- 12) Algoritma Divide and Conquer
- 13) Pendekatan teknik Greedy dan Dynamic programming
- 14) 10 Bahasa pemrograman populer diindonesia

b. Materi apa saja dan fungsi dalam penerapan pemrograman tersebut dalam kehidupan sehari-hari dan penerapannya dalam perkembangan teknologi terkini.

JAWAB

- 1) Berpikir Komputasional
- 2) Pemrograman Fungsional
- 3) Pemrograman Prolog
- 4) Pendekatan berorientasi objek
- 5) Apa itu pemrograman even driven
- 6) Memahami paradigma pemrograman procedural
- 7) Algoritma Brute Force dalam pencarian String
- 8) Algoritma Divide and Conquer
- 9) Pendekatan teknik Greedy dan Dynamic programming

c. Berikan usulan penelitian berikutnya terkait materi-materi yang telah anda pelajari!

JAWAB

- Mengenai pentingnya, dan implementasi berpikir komputasi bagi siswa.
- Implementasi pemrograman bagi kehidupan sehari-hari.