

IMPLEMENTASI FUZZY LOGIC BASED RANKING FUNCTION PADA SISTEM INFORMASI PENCARIAN ARTIKEL PARIWISATA BERBASIS MOBILE

Marvin Zeson Abilo, Diana Purwitasari, Rizky Januar Akbar.

Jurusan Teknik Informatika, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

e-mail: diana@if.its.ac.id

Abstrak— Kebutuhan manusia terhadap perkembangan teknologi informasi semakin meningkat. Begitu pula dengan kebutuhan manusia untuk mendapatkan sebuah informasi dengan cepat. Sebuah informasi didapatkan dengan menggunakan sebuah sistem temu kembali informasi yang mengembalikan sejumlah dokumen relevan dari sebuah dataset yang besar sesuai dengan query masukan pengguna. Pengembalian informasi relevan ini dilakukan berdasarkan hasil pembobotan dokumen dengan menggunakan metode tertentu. Metode paling umum yang digunakan dalam pembobotan dokumen adalah pembobotan TF-IDF (*term frequency-inverse document frequency*) yang merupakan hasil perkalian *term frequency* dengan *inverse document frequency*.

Pada tugas akhir ini akan diimplementasikan sebuah metode pembobotan dokumen yang menggunakan pendekatan fuzzy untuk meningkatkan akurasi dari sebuah sistem temu kembali informasi. Metode tersebut akan menggunakan beberapa elemen yang sering digunakan dalam pembobotan dokumen, seperti *term frequency*, *inverse document frequency* dan *normalization*. Metode ini akan melakukan pendekatan fuzzy dengan 2 tingkat untuk menghitung nilai relevansi dari sebuah dokumen.

Hasil uji coba menunjukkan metode dengan pendekatan fuzzy menghasilkan nilai *precision* dan *recall* yang lebih baik jika dibandingkan dengan metode non fuzzy.

Kata Kunci—Fuzzy Logic, Mobile, Pariwisata, Pembobotan Dokumen, Sistem Temu Kembali Informasi.

I. PENDAHULUAN

TEKNOLOGI informasi yang terus berkembang dengan pesat, sistem temu kembali informasi menjadi salah satu bidang yang banyak menjadi fokus pengembangan. Sistem temu kembali informasi banyak digunakan untuk mencari informasi dari sekumpulan data yang banyak, seperti pencarian halaman *website*, pencarian artikel, pencarian *file*, dan sistem rekomendasi. Masalah terbesar yang ada dalam pengembangan sistem temu kembali informasi adalah bagaimana sebuah sistem dapat mengembalikan informasi sesuai dengan yang diinginkan oleh pengguna dari sebuah sekumpulan data yang berjumlah besar. Sistem temu kembali informasi menggunakan sebuah *ranking function* untuk melakukan perhitungan nilai relevansi antara sebuah dokumen dengan *query* masukan dari

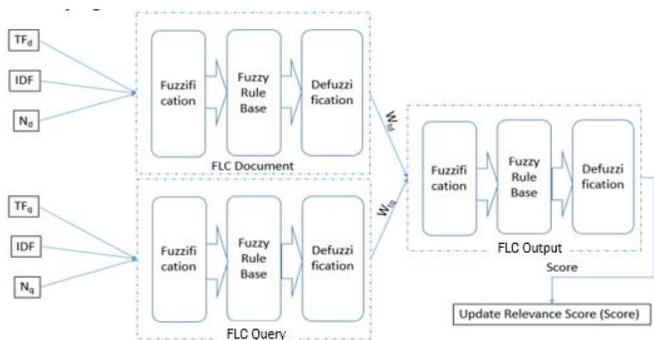
pengguna. Salah satu metode yang paling umum digunakan untuk melakukan perhitungan relevansi dokumen dengan *query* adalah metode TF-IDF yang merupakan hasil perkalian antara *term frequency* dengan *inverse document frequency* (Robertson, Walker, & Beaulieu, 1999).

Dalam tugas akhir ini akan diimplementasikan sebuah metode pembobotan baru yang menerapkan pendekatan fuzzy. Metode ini akan mengolah beberapa elemen yang sering digunakan dalam pembobotan dokumen, seperti *term frequency*, *inverse document frequency*, dan *normalization*. Metode fuzzy yang akan diimplementasikan merupakan pendekatan 2 tingkat yang dimana pada tingkat pertama terdapat 2 buah *fuzzy logic controller* yang masing-masing digunakan untuk mengolah dokumen dan mengolah *query*. Pada tingkat kedua terdapat 1 buah *fuzzy logic controller* yang digunakan untuk mengolah hasil keluaran *fuzzy logic controller* pada tingkat pertama.

Metode dengan pendekatan fuzzy ini akan diterapkan pada sebuah sistem temu kembali informasi yang berisi *data* berkategori pariwisata. *Categorical data* merupakan sebuah data yang hanya berisi *data* dengan kategori tertentu (Dictionary, n.d.). Dalam proses implementasi ini akan dilakukan beberapa *preprocessing data* seperti *stopword removal*, perhitungan *term frequency*, perhitungan *inverse document frequency*, dan perhitungan *normalization* untuk meningkatkan relevansi hasil pencarian (Rouse, 2005).

Metode pembobotan ini akan diimplementasikan dalam bentuk *web service* yang menerima *request* dan mengirimkan *response* kepada *client* yang berbentuk perangkat *mobile*. Perangkat *mobile* hanya digunakan untuk memasukkan *query* dan menampilkan hasil pencarian, sedangkan seluruh proses perhitungan akan dilakukan di *server*.

Bagian selanjutnya dari makalah ini memiliki struktur sebagai berikut. Pada bagian 2 akan dibahas tentang perancangan sistem dan algoritma yang akan diimplementasikan. Pada bagian 3 akan dibahas hasil uji coba yang telah dilakukan. Pada bagian 4 akan dibahas kesimpulan dan saran yang didapatkan dari hasil pengerjaan tugas akhir ini.



Gambar. 1. Struktur Metode Pembobotan Fuzzy 2 Tingkat

II. PERANCANGAN SISTEM

A. Metode Pembobotan

Metode pembobotan yang digunakan dalam tugas akhir ini didasarkan pada metode pembobotan *fuzzy* 2 tingkat (Gupta, Saini, & A.K., 2014). Pembobotan dengan pendekatan *fuzzy* 2 tingkat ini terdiri dari beberapa *fuzzy logic controller* yang digunakan untuk seluruh perhitungan *fuzzy* seperti *fuzzification*, operasi *rule fuzzy*, dan *defuzzification*. Pada tingkat pertama terdapat dua buah *fuzzy logic controller* yang masing-masing digunakan untuk perhitungan tingkat relevansi antara *term* dengan dokumen dan perhitungan seberapa pentingnya sebuah *term* di dalam dokumen. Tiap *fuzzy logic controller* pada tingkat pertama menerima tiga buah parameter yaitu *term frequency*, *inverse document frequency*, dan *normalization document length*. Pada tingkat kedua terdapat sebuah *fuzzy logic controller* yang digunakan untuk mengolah keluaran dari *fuzzy logic controller* yang ada pada tingkat pertama. Keluaran dari metode ini adalah nilai relevansi antara sebuah *term* dengan dokumen dalam skala 0 hingga 1. Gambaran dari metode pada penjelasan diatas ditunjukkan pada Gambar 1. Setiap notasi dan entitas yang ada pada gambar tersebut dijelaskan pada Tabel 1.

Sama seperti sistem *fuzzy* pada umumnya, proses pembobotan dokumen dengan menggunakan pendekatan *fuzzy* ini memiliki 3 buah proses yaitu *fuzzification*, operasi menggunakan *rule fuzzy*, dan *defuzzification*.

B. Fuzzification

Proses ini adalah proses pemetaan setiap parameter yang digunakan dalam sistem *fuzzy* kedalam nilai *fuzzy*. Setiap parameter yang masuk kedalam sistem *fuzzy* akan dipetakan dengan menggunakan *membership function* untuk mendapatkan nilai keanggotaan sebuah parameter pada tiap kelas *fuzzy* yang ada.

Membership function merupakan sebuah kurva yang mendefinisikan setiap nilai yang mungkin didapatkan dari sebuah parameter kedalam nilai keanggotaan dari sebuah kelas *fuzzy*. *Triangular type membership function* digunakan untuk menentukan nilai keanggotaan dari seluruh variabel dalam sistem *fuzzy*.

Range dari variabel masukan TF_d , TF_q , IDF , N_d , N_q direpresentasikan sebagai *Very Low* (VL), *Low* (L), *Medium* (M), *High* (H), *Very High* (VH). *Range* dari variabel W_{td} , W_{tq} ,

Tabel 1.
Deskripsi Notasi dan Entitas yang Digunakan dalam Metode Pembobotan dengan Pendekatan Fuzzy

Notasi	Deskripsi
TF_d	Frekuensi kemunculan suatu <i>term</i> didalam dokumen
IDF	Nilai <i>inverse document frequency</i> dari sebuah <i>term</i>
N_d	Nilai <i>inverse document length</i> dari sebuah dokumen
TF_q	Frekuensi kemunculan suatu <i>term</i> didalam <i>query</i>
N_q	Nilai <i>inverse document length</i> dari sebuah <i>query</i>
FLC Document	<i>Fuzzy logic controller</i> yang berfungsi untuk menghitung bobot sebuah <i>term</i> didalam dokumen
FLC Query	<i>Fuzzy logic controller</i> yang berfungsi untuk menghitung bobot sebuah <i>term</i> didalam <i>query</i>
W_{td}	Keluaran dari FLC Document
W_{tq}	Keluaran dari FLC Query
FLC Output	<i>Fuzzy logic controller</i> yang berfungsi untuk mengolah keluaran dari FLC Document dan FLC Query
Score	Keluaran dari FLC Output yang juga merupakan hasil pembobotan antara <i>term</i> dengan dokumen

dan Score direpresentasikan sebagai *Low* (L), *Medium* (M), *High* (H).

C. Tahap Perancangan Membership Function

Membership function dibuat berdasarkan *dataset* artikel yang tersedia. Proses penentuan *membership function* akan dilakukan sebagai berikut:

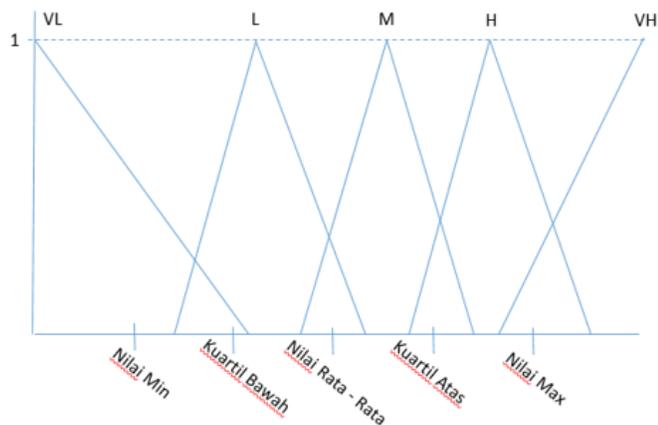
- 1) Analisa setiap parameter *fuzzy*.
- 2) Tentukan sampel *term* yang akan digunakan pada proses analisa dengan cara mencari *term* yang berkaitan dengan tema pariwisata dan memiliki nilai rata-rata *term frequency* yang tinggi.
- 3) Lakukan analisa dengan menentukan nilai tertinggi dan terendah yang bisa didapatkan oleh tiap variabel. Tentukan juga nilai rata-rata, kuartil atas, dan kuartil bawah.
- 4) Tentukan batas-batas tiap kelas pada *membership function* menggunakan nilai-nilai yang didapatkan dari proses analisa.

Contoh aturan yang digunakan untuk menentukan batas-batas kelas *membership function* berdasarkan nilai-nilai yang didapatkan dari proses analisa ditunjukkan pada Gambar 2.

D. Operasi Rule Fuzzy

Nilai-nilai yang didapatkan dari hasil proses *fuzzification* selanjutnya akan dioperasikan berdasarkan *rule-rule fuzzy* yang sudah ditentukan. *Rule fuzzy* merupakan suatu pernyataan kondisional yang menghubungkan informasi yang diberikan menjadi sebuah tindakan. Umumnya *rule fuzzy* memiliki struktur IF-THEN seperti yang ditunjukkan pada Persamaan 1.

$$IF\ tf\ Low\ AND\ idf\ Low\ THEN\ weight\ Low \quad (1)$$



Gambar. 2. Aturan Penentuan Batas-Batas Kelas dalam Membership Function

E. Penentuan Rule Fuzzy

Rule fuzzy yang dapat dibentuk dalam sistem fuzzy berjumlah 250 rule pada tingkat pertama dan 9 rule pada tingkat kedua. Rule-rule ini didapatkan dari hasil kombinasi seluruh representasi kelas dari tiap variabel dalam sistem fuzzy.

Penentuan rule-rule yang akan digunakan dilakukan dengan cara melakukan penggabungan rule-rule yang ada. Penggabungan ini akan dilakukan pada rule-rule yang memiliki tingkat kemiripan yang tinggi. Contoh proses penggabungan rule ini ditunjukkan pada Gambar 3.

Rule 1 = IF tf Low AND idf Low AND norm high THEN weight Medium

Rule 2 = IF tf Very Low AND idf Low AND norm very high THEN weight Medium

Rule Gabungan = IF tf Very Low OR Low AND idf Low AND norm very high OR high THEN weight Medium

Gambar. 3. Proses Penggabungan Rule Fuzzy

F. Defuzzification

Defuzzification adalah proses yang digunakan untuk mengubah fuzzy value yang merupakan hasil dari operasi rule fuzzy menjadi sebuah crisp value. Dalam metode pembobotan ini, proses defuzzification dilakukan dengan metode perhitungan centroid. Jika C_i adalah fuzzy value dari rule ke i dan Z_i adalah pemetaan fuzzy value dari rule i kedalam crisp value. Persamaan defuzzification dapat dituliskan pada Persamaan 2.

$$Defuzzification(C_i) = \frac{\sum_{i=0}^n C_i Z_i}{\sum_{i=0}^n C_i} \quad (2)$$

III. UJI COBA

A. Precision dan Recall

Uji coba dilakukan pada dataset artikel pariwisata yang diambil dari beberapa website. Berikut adalah beberapa website yang digunakan sebagai sumber pengambilan artikel:

1. Wisataindonesia.co.id
2. Indonesiawisata.info

3. Initempatwisata.com
4. Anekatempatwisata.com

Seluruh artikel yang digunakan dalam uji coba merupakan artikel berbahasa Indonesia dan jumlah keseluruhan adalah 534 artikel. Performa dari metode pembobotan akan diukur berdasarkan nilai precision dan recall. Persamaan yang digunakan untuk menghitung precision dan recall ditunjukkan pada Persamaan 3 dan Persamaan 4.

$$Precision = \frac{|R_a|}{|A|} \quad (3)$$

$$Recall = \frac{|R_a|}{|R|} \quad (4)$$

Tiap variabel dalam persamaan diatas dapat dijelaskan sebagai berikut $|A|$ adalah jumlah dokumen yang berhasil diambil oleh metode pembobotan yang diujikan dan $|R|$ adalah jumlah seluruh dokumen yang relevan dengan masing-masing query uji coba, dan $|R_a|$ adalah jumlah dokumen relevan yang berhasil diambil oleh metode pembobotan.

Tabel 2. Hasil Pengujian Metode Pembobotan dengan Pendekatan Fuzzy

Metode	10 Dokumen Pertama		20 Dokumen Pertama	
	Precision	Recall	Precision	Recall
Pendekatan fuzzy	0.6182	0.3926	0.5000	0.6361
Tanpa penggunaan rule gabungan	0.5909	0.3777	0.5045	0.6424
Tanpa parameter norm	0.6091	0.3966	0.4636	0.6083
Tanpa FLC query	0.6091	0.3809	0.4727	0.6247
Tanpa fuzzy	0.3273	0.2116	0.3091	0.4008

B. Hasil Uji Coba

Uji coba dilakukan dengan beberapa skenario yaitu:

1. Uji coba pembobotan dengan pendekatan fuzzy yang menggunakan seluruh parameter dan rule hasil penggabungan.
2. Uji coba pembobotan dengan pendekatan fuzzy yang menggunakan seluruh parameter tapi tidak menggunakan rule hasil penggabungan.
3. Uji coba pembobotan dengan pendekatan fuzzy tanpa menggunakan parameter norm.
4. Uji coba pembobotan dengan pendekatan fuzzy tanpa menggunakan FLC query.
5. Uji coba pembobotan TF-IDF standar.

Dari hasil uji coba yang ditunjukkan pada Tabel 2, ditunjukkan bahwa tidak terdapat perbedaan yang signifikan antara setiap pembobotan yang menggunakan pendekatan fuzzy. Perbedaan yang signifikan terlihat jika dilakukan perbandingan antara pembobotan yang menggunakan

pendekatan *fuzzy* dengan pembobotan yang tidak menggunakan pendekatan *fuzzy* (TF-IDF standar).

IV. KESIMPULAN

Tugas akhir ini memiliki beberapa kesimpulan yang dapat diambil berdasarkan data-data yang didapatkan pada bagian uji coba. Kesimpulan yang dapat diambil adalah sebagai berikut:

1. Pembobotan artikel dengan menggunakan pendekatan *fuzzy* meningkatkan hasil akurasi pencarian secara signifikan jika dibandingkan dengan metode *non fuzzy*.
2. Penghapusan parameter *norm* tidak memberikan perubahan yang signifikan terhadap akurasi hasil pencarian.
3. Penggunaan FLC *query* tidak terlalu berpengaruh terhadap akurasi hasil pencarian.
4. Penggunaan seluruh *rule fuzzy* dan *rule* yang sudah digabungkan memberikan hasil yang relatif sama.

DAFTAR PUSTAKA

- Dictionary, B. (n.d.). *What is categorical data? Definition and meaning*. Retrieved from Business Dictionary:
<http://www.businessdictionary.com/definition/categorical-data.html>
- Gupta, Y., Saini, A., & A.K., S. (2014). A new fuzzy logic based ranking function for efficient Information. *Expert System with Applications*, 1223-1234.
- Robertson, S., Walker, S., & Beaulieu, M. (1999). Okapi-BM25 at TREC-7: Automatic ad hoc, filtering, VLC and filtering tracks. *Proceedings of the seventh text retrieval conference*.
- Rouse, M. (2005, September). *What is data preprocessing?* Retrieved from Techtargget:
<http://searchsqlserver.techtarget.com/definition/data-preprocessing>