NAMA : NANDA PRAYOGI

NIM : A11.2021.13505

MATA KULIAH : SISTEM TEMU KEMBALI INFORMASI

KELOMPOK : A11.4703

STATUS : PRAKTIKUM EVALUASI

PERBANDINGAN ANTAR METODE DALAM MENGEVALUASI KONSEP

SISTEM TEMU KEMBALI INFORMASI

ABSTRAK

Sistem Temu Kembali Informasi adalah bidang penting dalam ilmu computer yang tersu berkembang. Evaluasi konsep STKI merupakan aspek kunci dalam mengukur kinerja sistem ini. Dalam artikel ini, saya melakukan tinjauan literatur terhadap berbagai metode evaluasi konsep STKI yang telah diusulkan oleh peneliti dalam beberapa tahun terakhir. Kami menganalisis dan membandingkan beberapa metode tersebut untuk memberikan wawasan yang lebih mendalam tentang pendekatan yang paling efektif dalam mengevaluasi konsep STKI. Penelitian ini bertujuan untuk membantu peneliti dan praktisi dalam memilih metode evaluasi yang paling sesuai dengan kebutuhan mereka dalam mengembangkan sistem temu Kembali informasi yang canggih.

1. PENDAHULUAN

Dalam era informs digital yang semakin berkembang, STKI telah menjadi elemen kunci dalam pengelolaan data, pencarian informasi, dan pengambilan keputusan. STKI adalah teknologi yang memungkingkan pengguna untuk mencari dan mengambil informasi yang relevan dari berbagai sumber data, termasuk dokumen teks, multimedia, dan data terstruktur. Keberhasilan dan efektivitas STKI sangat tergantung pada evaluasi yang tepat dan metode evaluasi konsep STKI yang digunakan.

Dalam dunia akademik dan industry, banyak upaya telah dilakukan untuk mengembangkan dan meningkatkan sistem STKI. Sejumlah metode evaluasi konsep STKI diajukan untuk mengukur kemampuan kinerja suatu sistem. Evaluasi konsep STKI tidak hanya penting untuk mengukur kemampuan sistem dalam menemukan informasi yang relevan, tetapi juga untuk memahami kekurangan dan perbaikan yang mungkin diperlukan.

1. METODE EVALUASI KONSEP STKI

Metode evaluasi yang digunakan dalam penelitian ini adalah pendekatan pengukuran kinerja berdasarkan relevansi. Metode ini didasarkan pada prinsip dasar bahwa sistem STKI harus memberikan hasil yang paling relevan terhadap permintaan pengguna. Evaluasi relevansi adalah salah satu aspek kritis dalam mengukur efektivitas sistem temu kembali informasi.

* 1. Pemilihan Koleksi Dokumen

Untuk melakukan evaluasi yang tepat, diperlukan koleksi dokumen yang sesuai dengan topik atau domain tertentu. Dalam kasus ini akan menggunakan koleksi dokumen yang mencakup sejumlah besar sumber informasi terkait dengan STKI.

* 1. Pembentukan Kueri

Kueri atau permintaan adalah elemen kunci dalam evaluasi relevansi. Kueri yang mewakili kebutuhan pengguna akan dibuat berdasarkan contoh permintaan pengguna sebelumnya atau skenario penggunaan yang diidentifikasi.

* 1. Pengujian Sistem

Sistem STKI akan diuji menggunakan kueri yang dibuat dalam langkah sebelumnya. Hasil yang dihasilkan oleh sistem untuk setiap kueri akan direkam.

* 1. Penilaian Relevansi

Panel penilai akan menilai tingkat relevansi setiap hasil sistem terhadap kueri yang sesuai. Penilaian ini dapar melibatkan penilai manusia yang berpengalaman dalam domain yang relevan.

* 1. Perhitungan Kinerja

Kinerja sistem akan dihitung berdasarkan hasil penilaian relevansi. Metrik umum yang digunakan termasuk Precision, Recall, F-Measure, dan Mean Average Precision (MAP).

* 1. Analisis Hasil

Hasil evaluasi akan dianalisis untuk memahasi kekuatan dan kelemahan sistem STKI yang diuji. Ini akan membantu dalam membandingkan kinerja sistem dengan metode evaluasi lainnya.

1. PERBANDINGAN METODE EVALUASI KONSEP STKI

Dlam bagian ini, akan memeriksa dan menganalisis beberapa metode evaluasi konsep STKI yang telah diusulkan dalam literatur terkait.

* 1. Metode Evaluasi Precision

Precision adalah metrik yang mengukur sejauh mana hasil temu kembali sistem STKI relevan dengan kebutuhan pengguna. Ini fokus pada seberapa banyak item yang ditemukan adalah benar-benar relevan.

Precision dihitung dengan membagi jumlah dokumen yang relevan yang ditemukan oleh sistem dengan jumlah total dokumen yang ditemukan.

Precision =

* 1. Metode Evaluasi Recall

Recall mengukur sejauh mana sistem STKI dapat menemukan semua item yang relevan dari koleksi. Ini fokus pada kemampuan sistem untuk menemukan semua dokumen yang seharusnya ditemukan.

Recall dihitung dengan membagi jumlah dokumen relevan yang ditemukan oleh sistem dengan jumlah total dokumen yang seharusnya ditemukan.

Recall =

* 1. Metode Evaluasi F-Measure

F-Measure adalah metode yang menggabungkan precision dan recall menjadi satu metrik tunggal untuk mengukur kinerja sistem STKI. Ini memberikan gambaran holistik tentang kualitas sistem temu kembali dengan mengimbangi precision dan recall.

F-Measure dihitung dengan menggunakan rumus harmonic mean dari precision dan recall.

F-Measure =

* 1. Metode Evaluasi Mean Average Precision

Mean Average Precision digunakan untuk mengukur kualitas peringkat sistem STKI dengan mempertimbangkan presisi pada setiap titik peringkat. Semakin tinggi item relevan muncul dalam daftar peringkat, semakin tinggi nilai MAP.

Langkah-langkah dalam menghitung MAP yang pertama hitung precision pada setiap titik peringkat. Lalu kalikan precision dengan nilai biner relevansi (1 jika relevan, 0 jika tidak) pada setiap posisi. Setelah itu jumlahkan semua precision pada posisi yang relevan. Dan terakhir bagi jumlah tersebut dengan jumlah total item relevan.

MAP =

* 1. Metode Evaluasi Normalized Discounted Cumulative Gain

Normalized Discounted Cumulative Gain mengukur kualitas peringkat sistem STKI dengan memberikan penekanan pada peringkat yang lebih tinggi dalam hasil temu kembali. Ini mengukur relevansi item dan bobot posisi dalam peringkat.

NDCG dihitung dengan menghitung DCG (Discounted Cumulative Gain) terlebih dahulu dan kemudian mengonversinya menjadi nilai terstandarisasi dengan menghitung IDCG (Ideal DCG).

NDCG =

* 1. Kesimpulan Perbandingan

Precision mengukur akurasi temu kembali pada hasil. Recall mengukur sejauh mana sistem dapat menemukan semua item yang relevan. F\_Measure mengkombinasikan precision dan recall menjadi satu metrik yang memberikan gambaran holistik tentang kinerja sistem. MAP lebih cocok untuk mengevaluasi kualitas peringkat dan urutan. NDCG memberikan penekanan pada peringkat yang lebih tinggi dalam hasil temu kembali.

Ini akan memberikan wawasan tentang cara masing-masing metode dihitung dan kapan mereka paling sesuai dalam mengevaluasi sistem STKI. Dan juga mencatat bahwa dlam beberapa situasi, kombinasi metode evaluasi mungkin diperlukan untuk memberikan gambaran yang lebih lengkap tentnag kinerja sistem.

1. KESIMPULAN

Dalam rangkuman ini, dapat menyimpulkan bahwa evaluasi konsep STKI merupakan langkah yang krusial dalam pengembangan sistem temu kembali informasi yang efektif dan efisien. Dalam melakukan tinjauan literatur mendalam terhadap berbagai metode evaluasi yang ada dan menganalisis serta membandingkan beberapa di antaranya untuk memberikan pandangan yang lebih jelas.

Dari hasil pencarian mengenai informasi tersebut, beberapa temuan penting muncul. Pertama, melihat bahwa metode evaluasi yang didasarkan pada pengukuran akurasi terbukti efektif dalam situasi di mana ketepatan informasi yang ditemukan oleh STKI sangat penting, seperti dalam konteks sistem kesehatan atau penelitian ilmiah. Namun, metode ini dapat menjadi kurang relevan dalam situasi di mana aspek lain seperti kecepatan dan skalabilitas lebih diutamakan.

Kedua, kecepatan juga menjadi faktor penting dalam evaluasi konsep STKI, terutama dalam kasus aplikasi yang membutuhkan respon instan. Dalam situasi seperti mesin pencari web, perbedaan kecepatan eksekusi metode evaluasi dapat memiliki dampak besar pada pengalaman pengguna.

Ketiga, kompleksitas metode evaluasi juga perlu dipertimbangkan. Metode evaluasi yang rumit seringkali memerlukan sumber daya yang lebih besar dalam hal komputasi dan waktu, tetapi dapat memberikan hasil yang lebih akurat. Namun, metode yang lebih sederhana mungkin lebih praktis dalam beberapa kasus.

Selain itu, penting untuk mencatat bahwa pilihan metode evaluasi harus disesuaikan dengan kebutuhan spesifik dari aplikasi STKI yang sedang dikembangkan. Setiap metode memiliki kelebihan dan kekurangan masing-masing, dan peneliti atau praktisi harus mempertimbangkan faktor-faktor ini saat memilih metode yang sesuai untuk proyek mereka.

1. DAFTAR REFERENSI

<https://www.academia.edu/6899398/Bab_1_KONSEP_DASAR_SISTEM_TEMU_KEMBALI_INFORMASI>

<https://journal3.uin-alauddin.ac.id/index.php/khizanah-al-hikmah/article/view/3044/3526>

<https://ichi.pro/id/akurasi-presisi-recall-dan-skor-f1-108847525674158#google_vignette>

<https://machinelearningmastery.com/precision-recall-and-f-measure-for-imbalanced-classification/>