KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)
Terakreditasi Nomor 204/E/KPT/2022 | Vol. 5, No. 1, Januari (2024), pp. 269-280

Peran Algoritma *Stemming Nazief Adriani* Dalam Peningkatan Relevansi Pencarian Dokumen

Dewi Soyusiawaty[1], Az-Zahra[2]
[1,2]Universitas Ahmad Dahlan, Yogyakarta, Indonesia
E-mail: dewi.soyusiawaty@tif.uad.ac.id[1],
azzahra1800018244@webmail.uad.ac.id[2]

Abstract

Nazief Andriani's algorithm is a Stemming Algorithm in text-preprocessing as a support in improving Information Retrieval (IR) performance and the process of determining the similarity value of text documents. But in reality, there are still many Information Retrieval systems that do not meet user needs, where to display search results, documents can only be found if the user enters keywords that must be exactly the same or have the same words as the query. The aim of this research is to create a system to improve and recognize keyword variations in the search relevance of thesis documents to meet user needs which will make it easier to search for document titles. The method used in this research was to collect data using 2045 thesis title documents. The method used is Nazief Adriani's Stemming Algorithm to make it easier to categorize document titles with more varied search results. So, for this research stage, a website will be built to increase the relevance of the accuracy of the role of stemming in document searches with research stages including data collection, needs analysis, system design, system implementation and testing, system testing. This system can display document search keywords with varying affix results with Precision test results of 81.7% which shows the quality of how useful this document search system is. and a recall value of 100% which represents the quality of how complete the relevant results are displayed by the search system. With word processing research, searches for thesis document collections will be able to be managed well and improve document search performance which is more varied according to the needs of Informatics students as system users.

Keywords: *Nazief Andriani Algorithm; Information Retrieval (IR); Stemming*

Abstrak

Algoritma Nazief Andriani merupakan Algoritma Stemming pada text-preprocessing sebagai pendukung dalam meningkatkan performa Information Retrieval (IR) dan proses penentuan dalam nilai kemiripan dokumen teks. Tetapi pada kenyataanya masih banyak sistem Information Retrieval yang belum memenuhi kebutuhan pengguna dimana untuk menampilkan hasil pencarian dokumen hanya dapat ditemukan jika pengguna memasukkan kata kunci yang harus sama persis atau yang mempunyai kata yang sama dengan query. Tujuan penelitian ini membuat sistem untuk meningkatkan dan mengenali variasi kata kunci dalam relevansi pencarian dokumen skripsi untuk memenuhi kebutuhan pengguna yang akan memudahkan pencarian judul dokumen. Metode yang digunakan dalam penelitian ini, untuk pengumpulan datanya menggunakan 2045 dokumen judul skripsi. Untuk metode yang digunakan yaitu Algoritma Stemming Nazief Adriani untuk memudahkan pengkategorian judul dokumen dengan hasil pencarian yang lebih bervariasi. Maka untuk tahap penelitian ini akan dibangun Website untuk meningkatkan relevansi keakuratan peran stemming pada pencarian dokumen dengan tahapan penelitian meliputi pengumpulan data, analisis kebutuhan, perancangan sistem, implementasi dan pengujian sistem, pengujian sistem. Sistem ini dapat menampilkan sesuai kata kunci pencarian dokumen dengan hasil imbuhan yang bervariasi dengan hasil pengujian Precision sebesar 81,7% yang mempresentasikan kualitas seberapa berguna

ISSN: 2720-992X

sistem pencarian dokumen ini. dan nilai recall sebesar 100% yang mempresentasikan kualitas seberapa lengkap hasil relevan yang ditampilkan oleh sistem pencarian. Dengan adanya penelitian pemrosesan kata pencarian kumpulan dokumen skripsi akan dapat dikelola secara baik dan meningkatkan kinerja pencarian dokumen yang lebih bervariasi sesuai dengan kebutuhan mahasiswa Informatika sebagai pengguna sistem.

Kata Kunci: *Algoritma Nazief Andriani; Information Retrieval (IR); Stemming*

1. Pendahuluan

Algortima *Stemming* merupakan salah satu Algoritma yang tidak terpisahkan dalam meningkatkan performa *Information Retrieval (IR)*. istilah dari Pencarian informasi berupa dokumen atau teks biasa dikenal dengan *Information Retrieval (IR)* yaitu proses pemisahan dokumen - dokumen. yang dianggap relevan dari sekumpulan dokumen yang tersedia. *Stemming* adalah proses pemotongan (penghilangan) imbuhan *(affix)*, baik awalan (*prefix)* maupun *akhiran(suffix)*, dari sebuah *term* untuk mendapatkan kata dasar dari kata berimbuhan[5]. Beberapa algoritma *stemming* untuk Bahasa Indonesia telah dikembangkan sebelumnya antara lain Algoritma *Nazief & Adriani*, Algoritma *Porter*, Algoritma Algoritma Arifin Setiono[2]. Dari kesimpulan penelitian terdahulu bahwa Algoritma *Nazief Adriani* lebih unggul dalam hal kecepatan dan akurasi dibandingkan dengan dua Algortima lainnya. Algoritma *Nazief Adriani* dikembangkan berdasarkan aturan morfologi Bahasa Indonesia yang mengelompokkan imbuhan menjadi awalan, sisipan,akhiran dan gabungan awalan akhiran. Algoritma ini menggunakan kamus kata dasar dan mendukung *recoding*, yakni merangkai kembali kata-kata yang mengalami proses *stemming* berlebih. Untuk memeriksa apakah kata dasar yang melalui proses stemming benar dan ditemukan pada kamus saat proses *stemming* dilakukan maka dibutuhkan kamus kata dasar[12].

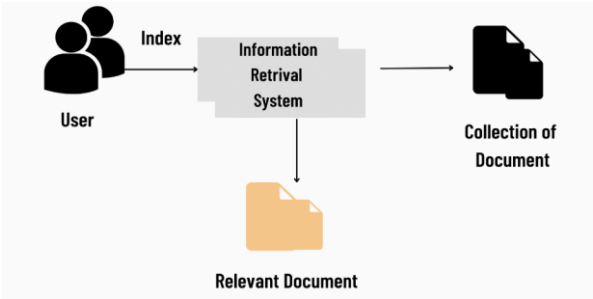
Pada umumnya indikator yang dipakai untuk menilai relevansi hasil pencarian suatu dokumen adalah menyesuaikan antara *query* yang diberikan dan dokumen yang diperoleh. Tetapi *term-term* yang terdapat pada dokumen dan pada *query* sering memiliki banyak varian morfologik, sehingga pasangan *term seperti "peningkatkan","meningkatkan"* dan *"tingkat"* tidak akan dianggap *ekuivalen* atau memiliki makna sama oleh sistem tanpa suatu bentuk *Natural Language Processing (NLP)[9]*. *Sistem Information Retrieval* yang ideal adalah dimana sistem dapat menemukan informasi yang relevan yang sesuai permintaan pengguna. Tetapi pada kenyataanya masih banyak *sistem Information Retrieval* yang belum memenuhi kebutuhan pengguna dimana untuk menampilkan hasil pencarian dokumen hanya dapat ditemukan jika pengguna memasukkan kata kunci yang harus sama persis atau yang mempunyai kata yang sama dengan *query*[11]. Seharusnya kata kunci dengan banyak variasi morfologi harus dipertimbangkan sebagai term yang sama. Dokumen yang memiliki kata yang merupakan variasi dari kata pada query tidak dianggap sebagai dokumen hasil pencarian. Tentu ini menjadi permasalahan karena sulit untuk mendapatkan hasil dan mengambil informasi yang relevan karena bahasa pastinya memiliki berbagai varian morfologi kata - kata yang akan mengakibatkan terjadinya ketidaksesuaian kosakata[9].

Maka pada penelitian ini akan dilakukan proses *stemming* Untuk mengenali variasi morfologi tersebut dalam meningkatkan relevansi pencarian dokumen dengan menggunakan Dokumen tugas akhir mahasiswa. Berdasarkan uraian di atas maka akan dibangun *Information Retrieval System* pada Proses Pencarian Dokumen Digital Menggunakan Metode Algoritma S*temming Nazief & Adriani* untuk meningkatkan relevansi pencarian dokumen dengan menggunakan *Library Sastrawi*. Dengan adanya penelitian pemrosesan kata ini diharapkan sistem pencarian dokumen akan memenuhi kebutuhan pengguna dalam meningkatkan dan mengoptimalkan relevansi pencarian dokumen untuk meningkatkan kinerja pencarian dokumen yang lebih bervariasi.

2. Metodologi Penelitian

2.1. *Information Retrieval* (IR)

Invormation Retrieval (IR) adalah proses pemisahan dokumen yang dianggap relevan dari sekumpulan dokumen yang tersedia. Gambar 1 merupakan *representasi arsitektur Information Retrieval (IR)* yang menggambarkan pencarian informasi[11].



Gambar 1. Arsitektur *Information Retrival (IR)*

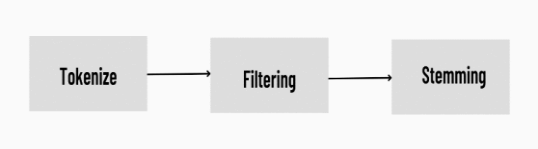
Penyajian dan Implementasi hasil informasi harus memadai sehingga pengguna dapat dengan mudah mendapatkan informasi yang diinginkan. Sistem pencarian dokumen harus efektif dan efisien. Maksud dari efektif yaitu efektif yaitu pengguna mendapatkan dokumen yang relevan dengan *query* yang di*input*kan dan efisien yaitu meningkatkan waktu pencarian yang digunakan sesingkat mungkin. *Search Engine* merupakan salah satu aplikasi *Invormation Retrieval* yang tak terlepas dari kebutuhan dari dukungan perangkat *indexing* dan *query expansion*. Proses untuk membangun *Information Retrieval System* pada Proses Pencarian Dokumen Digital yaitu dokumen disimpan, kemudian dokumen tersebut dikategorikan sesuai isi dari dokumen yang diperoleh secara otomatis dengan metode pencarian dokumen *text preprrocessing*. Dengan melakukan *preprocessing text* kata yang terdapat dalam dokumen akan dipilih dan ditetapkan kata mana yang dibutuhkan. Kata-kata yang terdapat dalam dokumen dan Informasi penting tersebut akan diambil untuk disimpan sebagai *tags* atau *keyword* dalam *database* yang nantinya untuk dicocokan dengan kata kunci pencarian. Kemudian saat pengguna memasukkan kata kunci pencarian maka hal ini akan diproses untuk mencari dokumen yang relevan. pada dasarnya *Information retrieval* merupakan proses untuk menentukan dokumen dalam koleksi yang harus dikelola dan diproses untuk memenuhi keinginan pengguna akan informasi.

Suatu sistem pencarian dokumen dikatakan ideal jika memenuhi kebutuhan pengguna dan menemukan semua dokumen yang relevan. Banyaknya varian morfologik pada suatu kata pada bahasa sangat berpengaruh terhadap hal tersebut, sehingga pemilihan algoritma *stemming* yang tepat sangat menentukan performa sistem *IR* juga proses *indexing* dan *query expansion* pada *Search Engine.* maka dari itu setiap kata harus diubah kedalam bentuk dasarnya atau dilakukan proses *stemming* untuk penyederhanaan bentuk dari suatu kata, sehingga sistem proses pengolaan kata menjadi efektif.

2.2. *Text Preprocessing*

Text processing merupakan suatu proses awal mengubah bentuk data dari yang belum terstruktur menjadi data yang terstruktur yang akan diolah dan di sesuaikan dengan kebutuhan[2]. Di dalam dokumen terdapat beragam kata, tetapi tidak semua kata yang ada di dalam dokumen itu dapat mewakilkan dokumen tersebut. Dengan melakukan *text preprocessing* kata yang terdapat dalam dokumen akan dipilih dan ditetapkan kata mana yang dibutuhkan. *Preprocessing* bertujuan mengubah teks menjadi *term index*, mengurangi *volume* kosa kata, menyeragamkan kata, dan menurunkan *noise* yang terdapat dalam sebuah dokumen.

Terdapat tahap-tahap dalam melakukan *text preprocessing*, yaitu : *Tokenisasi, Filtering, dan Stemming*. *Tokenisasi* adalah proses pemisahan kata yang berada dalam suatu kalimat,paragraf atau dokumen. *Filtering* adalah proses memisahkan kata yang dianggap tidak penting atau tidak memiliki makna. Dan *Stemming* adalah proses merubah kata yang berimbuhan menjadi kata dasar[3]. Berikut pada Gambar 2 proses tahapan *preprocessing[12]*.



Gambar 2. Tahapan *Preprocessing*

1) *Tokenize*: proses pemisahan kata yang berada dalam suatu kalimat,paragraf atau dokumen. Pada proses preprocessing untuk proses tokenisasi semua term pada dokumen yang akan di olah menjadi kecil (lower case) terlebih dahulu. Selain kata, tokenize juga menghilangkan angka, tanda baca, karakter karena dianggap tidak penting dalam pemrosesan kata. Tokenize memotong kumpulan karakter menjadi kata tunggal atau token[4].
2) *Filtering*: Pada tahap filtering ini merupakan tahap mengambil kata kata penting dari hasil tokenisasi yang dilakukan sebelumnya. Tahap filtering adalah proses memisahkan kata yang dianggap tidak penting atau tidak memiliki makna. Proses pemilihan kata ini dengan cara memilih kata yang penting dan membuang kata sambung atau kata yang tidak diperlukan dalam mewakili sebuah dokumen dan Kata sambung yang dibuang misalnya seperti yang, atau, maupun, untuk, agar, supaya dan kata sambung lainnya[4]. Filtering memproses kata hasil dari proses tokenisasi menjadi lebih sedikit dengan mengurangi kata yang termasuk ke dalam stopwords. Yang dimana pengeleminasian stopwords mempunyai keuntungan yaitu mengurangi jumlah space pada tabel term index hingga 40% atau bisa lebih.
3) *Stemming:* Proses setelah melakukan Filtering kata,tahap selanjutnya yaitu proses Stemming. Stemming adalah proses merubah kata yang berimbuhan menjadi kata dasar (root). Proses stemming ini dilakukan dengan menghilangkan semua imbuhan (affixes) yang terdiri dari awalan (prefixes), sisipan (infixes), akhiran (suffixes) dan confixes[7]. Algoritma stemming pada bahasa satu dan yang lain sangatlah berbeda. Contohnya Pada Bahasa Inggris memiliki perbedaan morfologi dari Bahasa Indonesia hal ini dapat dilihat dari kerumitan/kompleksnya teks Bahasa Indonesia yang terdapat beragam variasi imbuhan yang harus diolah untuk mendapatkan kata dasar (root). Dalam mendukung efektivitas dalam pencarian sebuah dokumen, penerjemahan dokumen, dan pencarian dokumen teks, proses Stemming sangatlah berperan penting . stemming merupakan proses penjabaran berbagai variasi bentuk dari suatu kata menjadi kata dasarnya. Secara luas roses dalam sistem Information retrieval dengan proses stemming sudah digunakan dalam pencarian informasi sebagai peningkatan kualitas informasi yang didapatkan untuk mendapatkan hubungan antara varian kata yang satu dengan yang lainnya. Untuk contoh yaitu kata “layanan”, “dilayani” dan “melayankan” yang awalnya memiliki makna yang berbeda dapat di stem menjadi sebuah kata “layan” yang memiliki arti yang sama sehingga kata-katanya saling berhubungan. Proses stemming juga dapat digunakan untuk mengurangi ukuran dari suatu ukuran index file. Contoh lain dalam suatu deskripsi terdapat varian kata “menggunakan”, “digunakan”, dan “pengguna” dan hanya memiliki kata dasar yaitu “guna”. Proses stemming pada teks Bahasa Indonesia memiliki perbedaan dengan stemming teks Bahasa Inggris karena pada teks bahasa Inggris prosesnya hanya menghilangkan sufiks (akhiran)[8]. Sedangkan pada teks Bahasa Indonesia, selain sufiks, prefiks, dan konfiks juga dihilangkan.

2.3. Algoritma *Stemming Nazief-Adriani*

Algoritma *Stemming Nazief-Adriani* adalah metode yang digunakan untuk meningkatkan efisiensi *Information Retrieval* dengan cara mengubah kata-kata dalam sebuah dokumen menjadi kata dasarnya[6]. Proses stemming teks Bahasa Indonesia digunakan untuk menghilangkan *akhiran, imbuhan, dan awalan.* Ini berbeda dari teks Bahasa Inggris karena proses stemming hanya digunakan untuk menghilangkan *sufiks*. Algoritma yang dibuat oleh *Bobby Nazief* dan *Mirna Adriani* memiliki tahapan sebagai berikut[7]:

1) Mencari kata yang akan distemming didalam kamus. Jika ditemukan maka diartikan kata tesebut adalah *k*ata dasar dan proses algoritma berhenti. Tetapi jika tidak ditemukan di lanjutkan ke proses tahapan selanjutnya.
2) Menghilangkan *Inflection Suffixes* (“kah”, “lah”, “ku”, “mu” atau “nya”). Jika *partikel* (“lah”, “kah”, “tah” atau“pun”) maka proses akan di ulangi lagi untuk menghapus Possesive Pronouns(“ku”, “ mu”, atau “-nya”), jika ada.
3) Proses selanjutnya menghilangkan *Derivation Suffixes* (“i”, “an” atau “kan”). Jika kata ditemukan pada kamus proses algoritma berhenti. Jika tidak maka ke tahap 3a.
    a. Jika “an” telah dihapus dan huruf terakhir dari kata tersebut adalah “- k”, maka “-k” juga akan ikut dihapus. Jika kata tersebut ditemukan dalam kamus maka algoritma berhenti. Jika tidak ditemukan maka lakukan tahap 3b.
    b. Akhiran yang dihapus (“-i”, “-an” atau “-kan”) dikembalikan, lanjut ke tahap 4.
4) Menghapus *Derivation Prefix*. Jika pada tahap 3 ada *sufiks* yang dihapus maka pergi ke tahap 4a, jika tidak pergi ke tahap 4b.
    a. Periksa tabel kombinasi awalan akhiran yang tidak diizinkan. Jika ditemukan maka algoritma berhenti, jika tidak pergi ke tahap 4b.
    b. For i = 1 to 3, tentukan tipe awalan kemudian hapus awalan. Jika kata dasar belum juga ditemukan lakukan tahap 5, jika sudah maka algoritma berhenti. Catatan: jika awalan kedua sama dengan awalan pertama algoritma berhenti.
5) Melakukan Recoding.
6) Jika semua tahapan telah selesai tetapi tidak juga berhasil maka kata awal di asumsikan sebagai kata dasar. Proses selesai.

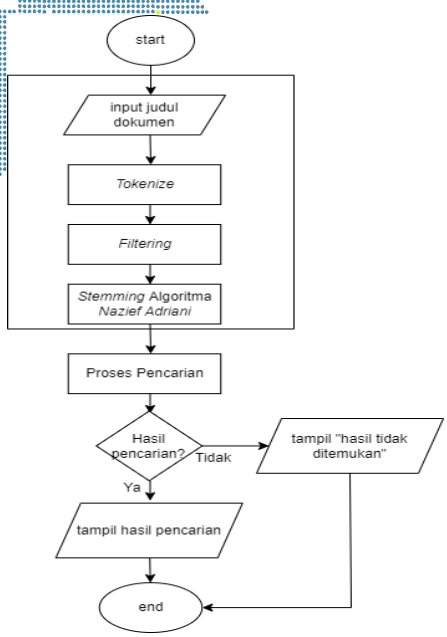
3. Hasil dan Pembahasan

3.1. Analisis Kebutuhan Data

Data yang dikumpulkan merupakan data judul skripsi mahasiswa dengan total 2046 data judul dokumen. Data masing-masing meliputi Id dokumen,judul dokumen,penulis dan tahun terbit.

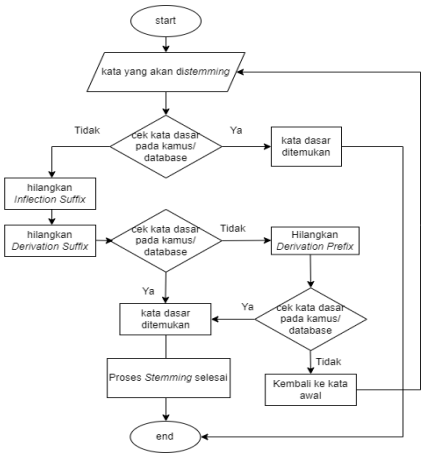
3.2. Perancangan Sistem

Setelah melakukan proses tahapan penelitian, tahap selanjutnya yaitu melakukan perancangan sistem. Pada tahap ini dilakukan pengimplementasi supaya sistem dapat berjalan dengan baik. Alur perancangan sistem dapat dilihat pada Gambar 3.



Gambar 3. Perancangan Sistem

Berikut diagram alir Algoritma *Stemming Nazief Adriani* pada sistem pencarian dokumen pada Gambar 4.



Gambar 4. Diagram Alir Algoritma *Stemming Nazief Adriani*

Pada proses Diagram alir Algoritma *Stemming Nazief Adriani* dilakukan untuk mencari kata dasar sebagai berikut :

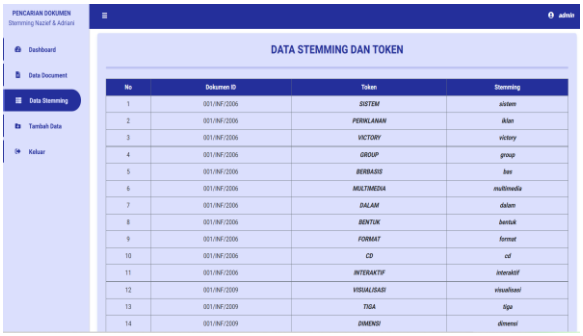
a. Proses ini diawali dengan Kata yang akan di stemming dicari pada kamus atau database jika ditemukan maka proses algoritma selesai.
b. Jika kata dasar tidak ditemukan maka lanjut ke tahap selanjutnya yaitu menghilangkan *inflection suffix*(awalan).
c. dilanjuti menghilangkan Derivation *Suffixs*(akhiran).
d. Kemudian cek Kembali kata kunci pada kamus atau *database* jika kata dasar ditemukan maka proses selesai dan proses berhenti.
e. Jika tidak ditemukan maka ketahap proses menghilangkan *Derivation Prefix* dan proses algortima berhenti jika kata dasar ditemukan pada kamus atau database.
f. Jika tidak ditemukan maka lakukan *recoding*

3.3. Implementasi Tampilan

1) Data Dokumen: Pada menu data dokumen ini berisi informasi dari dokumen yaitu id dokumen judul dokumen penulis dan tahun dokumen. dan proses *stemming* dapat dilakukan pada semua data. Tampilan dapat dilihat pada Gambar 5.

Gambar 5. Data Dokumen

2) Data *Stemming:* Pada menu ini dapat dilakukan proses *stemming* semua judul dokumen. Tampilan dapat dilihat pada Gambar 6.



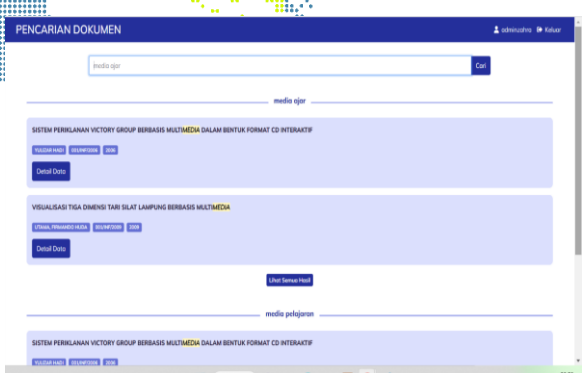
Gambar 6. Tampilan data *Stemming*

3) Halaman Pencarian Dokumen Pengguna: Pada menu tampilan ini pencarian dokumen dimana pengguna memasukkan kata kunci judul dokumen yang akan dilakukan proses pencarian dengan kata kunci “media ajar” .berikut tampilan pada Gambar 7.

Gambar 7. Halaman Pencarian Dokumen *User*

4) Variasi Kata Kunci: Pada menu proses pencarian judul dokumen dan sistem akan menampilkan hasil pencarian dari kata kunci “media ajar” dan menampilkan hasil imbuhan variasi dari kata kunci yang dimasukkan

yaitu "media ajar" dan "media pelajaran" .Dapat dilihat tampilan dari pencarian dokumen pada Gambar 8.



Gambar 8. variasi pencarian kata kunci media ajar

Pada tampilan ini hasil variasi dari imbuhan "media ajar" dengan 376 hasil judul dokumen yang memiliki unsur kata "media ajar". berikut pada Gambar 9.

Gambar 9. Hasil pencarian dokumen kata media ajar.

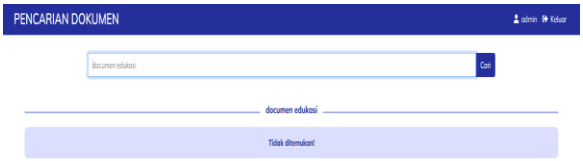
Kemudian pada tampilan ini hasil variasi dari imbuhan "media pelajaran" dengan 315 hasil judul dokumen yang memiliki unsur kata "media pelajaran". berikut pada Gambar 10.

Gambar 10. Hasil pencarian dokumen kata Media Pembelajaran

Pada menu pencarian dokumen user juga dapat menampilkan detail data yaitu nama penulis id dokumen judul dan kata stemming. Berikut menu tampilan pada Gambar 11.

Gambar 11. Menu tampilan detail data

Pada Gambar 12 merupakan tampilan jika kata kunci judul dokumen yang dimasukkan oleh user tidak terdapat pada sistem database atau tidak ditemukan.



Gambar 12. judul dokumen tidak ditemukan

3.4. Pengujian Sistem

1) *Pengujian Recall & Precision:* Berikut pada Tabel 1 variasi kata kunci yang akan diuji terdiri dari 12 variasi pengujian dengan kata kunci yang berhubungan dengan judul dokumen informatika.

Tabel 1. Variasi Kunci Yang Diuji

No	Nama	Variasi kata kunci yang dicari
1	Pengujian 1	Naive Bayes
2	Pengujian 2	Layanan Informasi
3	Pengujian 3	Sistem Pendukung Keputusan
4	Pengujian 4	Multimedia
5	Pengujian 5	Peningkatan
6	Pengujian 6	Pengolahan Data
7	Pengujian 7	Perangkat Lunak
8	Pengujian 8	Perancangan Algoritma
9	Pengujian 9	Pengembangan
10	Pengujian 10	Media Ajar
11	Pengujian 11	Sistem Informasi
12	Pengujian 12	Berbasis Web

Selanjutnya pada Tabel 2 yaitu proses koleksi yang akan dicari.

Tabel 2. Koleksi Yang Dicari

No	Koleksi yang dicari	Ditemukan	Relevan	Tidak Relevan	Keterangan
1	Naive Bayes	15	15	0	Menemukan 15 dokumen semua relevan
2	Layanan Informasi	520	26	494	Dari 520 hasil pencarian

No	Koleksi yang dicari	Ditemukan	Relevan	Tidak Relevan	Keterangan
					Ditemukan 26 yang relevan & 494 yang tidak relevan
3	Sistem Pendukung Keputusan	150	150	0	Menemukan 150 dokumen semua relevan
4	Multimedia	200	200	0	Menemukan 200 dokumen semua relevan
5	Peningkatan	28	5	23	Dari 28 hasil pencarian ditemukan 5 yang relevan & 23 yang tidak relevan
6	Pengolahan Data	3	3	0	Menemukan 3 dokumen semua relevan
7	Perangkat Lunak	26	26	0	Menemukan 26 dokumen semua relevan
8	Perancangan Algoritma	132	98	34	132 hasil pencarian ditemukan 98 yang relevan &34 yang tidak relevan
9	Pengembangan	503	405	98	503 hasil pencarian ditemukan 405 yang relevan & 98 yang tidak relevan
10	Media Ajar	376	221	151	376 hasil pencarian ditemukan 221 yang relevan &151 yang tidak relevan
11	Sistem Informasi	1015	1015	0	Menemukan 1015 dokumen semua relevan
12	Berbasis Web	830	830	0	Menemukan 830 dokumen semua relevan

Tabel 3.. Variasi Kunci Yang Diuji

No	Kata Kunci	Hasil Pengujian Relevan	Hasil Pengujian Tidak Relevan
1	Layanan Informasi	Layanan Informasi	Pelayanan Informasi
			Informasi pelayanan
			Informasi layanan
			Pelayanan
2	Peningkatan	Peningkatan	Tingkat
			Tingkatan
			Bertingkat
3	Media Ajar	Media Ajar	Multimedia
		Media Pembelajaran	Media

Pada Tabel 3 contoh hasil pengujian untuk kata kunci “layanan informasi” dan “peningkatan” dikatakan relevan disini jika kata kunci yang ditemukan dari kata “layanan informasi” dan “peningkatan” tanpa tambahan imbuhan apapun. Sedangkan tidak relevan yaitu ditemukan imbuhan dari variasi kata kunci “layanan informasi”seperti “pelayanan

informasi", "informasi pelayanan". Dan jika hanya kata "informasi", "layanan" dan "pelayanan" yang tampil dihitung sebagai tidak relevan tetapi kata kunci tetap ditemukan. Begitu juga pada variasi kata kunci "peningkatan" seperti "tingkat", "tingkatan" dan "bertingkat". Untuk kata kunci "media ajar" dengan variasi "media Pembelajaran", "multimedia", "media", "ajar". Begitu juga pada kata kunci pengujian yang lain terdapat variasi dari kata kunci yang di tampilkan. Sedangkan jika semua kata kunci relevan tidak ada variasi dari kata kunci tersebut.

Tabel 4.. Hasil Uji

No	Relevan (a)	Tidak relevan (b)	Total (a+b)	Tidak ditemukan (c)	Total (a+c)	Recall |a/(a+c)| x100%	Precision |a/(a+b)| x100%
1	15	0	15	0	15	100	100
2	26	494	520	0	520	100	50.0
3	150	0	150	0	150	100	100
4	200	0	200	0	200	100	100
5	5	23	28	0	5	100	17.857143
6	3	0	3	0	3	100	100
7	26	0	26	0	26	100	100
8	98	34	132	0	98	100	74.242424
9	405	98	503	0	405	100	80.516899
10	221	151	376	0	221	100	58.776596
11	1015	0	1015	0	1015	100	100
12	830	0	830	0	830	100	100
	Rata Rata					100	81.7827552

Dari uraian Tabel 4 dari skala 0%-100% didapatkan hasil rata rata nilai *Precision* sebesar 81,7% yang mempresentasikan kualitas seberapa berguna sistem pencarian dokumen ini. dan nilai *recall* sebesar 100% yang mempresentasikan kualitas seberapa lengkap hasil relevan yang ditampilkan oleh sistem pencarian. Maka didapatkan pengujian nilai *precision* lebih rendah dari pada nilai *recall* berdasarkan kata kunci yang digunakan oleh pengguna sistem relevansi pencarian dokumen. Tetapi tingkat keefektifan dari sistem relevansi pencarian dokumen sudah dapat dikatakan efektif.

4. Kesimpulan

Sistem pencarian dokumen dengan metode Algoritma Stemming Nazief Adriani memiliki pembaharuan yang dapat mempermudah pengguna dalam proses pencarian judul dokumen dan menampilkan hasil referensi pencarian yang lebih banyak dan variative. Hasil pengujian Precision sebesar 81,7% yang mempresentasikan kualitas seberapa berguna sistem pencarian dokumen ini. dan nilai recall sebesar 100% yang mempresentasikan kualitas seberapa lengkap hasil relevan yang ditampilkan oleh sistem pencarian. Hasil pengujian System Usubility Scale (SUS) sebesar 78% dimana sistem sudah memenuhi kebutuhan pengguna. dan pengujian Blackbox Testing sebesar 93% dimana sistem berjalan dengan baik. Adapun beberapa masukan untuk pengembangan sistem relevansi pencarian dokumen selanjutnya yaitu diharapkan tidak hanya menggunakan Algoritma stemming tetapi dapat lebih banyak menggunakan algoritma lain yang bisa di kembangkan lagi.

Daftar Pustaka

[1] Sardjono, M,Wisuda., Cahyanti, Margi., Mujahidin, Maulana. Ariany, Rini., (2018). Pendeteksi Kesamaan Kata untuk judul penulisan Berbahasa Indonesia menggunakan Algoritma *Stemming* Nazief - Adriani. Sebatik. VOL 22 NO 2 (2018).

[2] Sari, Bunga., Sibaroni, Yuliant., (2019). Deteksi Kemiripan Dokumen Bahasa Indonesia Menggunakan Algoritma *Smith Waterman* dan Algoritma Nazief & Andriani. Indonesia Jurnal Of Computing. Vol. 4, Issue. 3, Dec 2019.
[3] Widiasri, M., Tjandra, E., Chandra, Lisa Maria., (2017) Peningkatan Kinerja Pencarian Dokumen Tugas Akhir Menggunakan Porter Stemmer Bahasa Indonesia dan Fungsi Peringkat Okapi BM25.TEKNIKA, Volume 6, Nomor 1.
[4] Wirayasa,Putu Merta., Wirawan, Made Agus.,(2019). Adaptasi Algoritma Nazief Adriani untuk Stemming Teks Bahasa Bali. Janapati UNDHISKA.Volume 8, Nomor 1, Maret 2019.
[5] Ahmad, Rahmat., Sasue, Riz Rifai Oktavianus., (2020. Sistem Penilaian Esai Otomatis Menggunakan Algoritma Stemming Nazief Adriani. Politeknik Transportasi Darat Bali. Vol. 1 No. 2 (2020).
[6] Pramudita, Hafiz Ridha., (2018). Penerapan Algoritma Stemming Nazief & Adriani Dan Similarity Pada Penerimaan Judul Thesis. Jurnal Ilmiah DASI.Vol. 15 No. 04 Desember 2018.
[7] Wahyudi, Dwi., Susyanto, Teguh., Nugroho,Didik.,(2017). Implementasi dan Analisis Algoritma Stemming Nazief & Adriani dan Porter pada Dokumen Berbahasa Indonesia. Jurnal Ilmiah SINUS. Vol 15, No 2 (2017).
[8] Manase, Sahat H Simarangkir., (2017). Studi Perbandingan Algoritma - Algoritma Stemming untuk Dokumen Teks Bahasa Indonesia. Jurnal Infokar. Volume 1 No. 1 2017.
[9] Magriyanti, Arie Atwa., (2018). Analisis Pengembangan Algoritma Porter Stemming dalam Bahasa Indonesia. Sekolah Tinggi Elektronika dan Komputer PAT. September 29, 2018.
[10] Martin., Nilawati, Lala., (2019). Recall dan Precision Pada Sistem Temu Kembali Informasi Online Public Access Catalogue (OPAC) di Perpustakaan. Jurnal Komputer dan Informatika Universitas Bina Sarana Informatika. Volume XXI No. 1 Maret 2019.
[11] Setiawan, Hamzahi., (2021). Implementasi Algoritma Fuzzy Topsis Pada Sistem Rekomendasi Beasiswa. Jurnal Sistem Telekomunikasi Elektronika Sistem Kontrol Power Sistem & KomputerVol.1 / No 2.13 Juli, 2021.
[12] Herawan,Agus., Hakim, Patria Rahman., Pamadi, Bambang Sigit.,(2020). Perangkat Lunak Search Engine Citra Satelit LAPAN-A2 dan LAPAN-A3. Jurnal Edukasi & Penelitian Informatika. Vol 6, No 2 (2020).
[13] Guerra, Cecilia., Costa, Nilza., (2020). Can Pedagogical Innovations Be Sustainable? One Evaluation Outlook for Research Developed inPortuguese Higher Education. Education science. 11 November 2021.
[14] Novitasari, Dian., (2016). Perbandingan Algoritma Stemming Porter dengan Arifin Setiono untuk Menentukan Tingkat Ketepatan Kata Dasar. Universitas Indraprasta PGRI. Vol 1, No 2 (2016).
[15] Pratama, Enda Esyudha., (2018). Information Retrieval pada Proses Penyimpanan dan Pencarian Dokumen Digital Menggunakan Metode Text Mining. Seminar Nasional Teknologi Informasi dan Komunikasi (SEMNASTIK). Vol 1 No 1 (2018).
[16] Ariyani, Pipin Farida., Rahmala, Annisa., Juliasari, Noni.,(2019). Implementasi Metode Stemming Tala Dan Fungsi Jaccard Pada Aplikasi Katalog Perpustakaan. Seminar Nasional Inovasi dan Aplikasi Teknologi di Industri 2019. ISSN 2085-421.