

PENERAPAN ALGORITMA FUZZY C-MEANS PADA PENGKLASTERAN HASIL PENCARIAN HADITS SHAHIH BERBASIS MOBILE

Rahadi Deli Saputra¹, Dani Rohpandi², Ihsan Jatnika Fahmillah³
^{1,2,3}Jurusan Teknik Informatika STMIK Tasikmalaya, Jalan RE. Martadinata 272A, 0265-
310830 Tasikmalaya
e-mail: rahadisianipar@gmail.com¹, danirtms@gmail.com², ihsanjatnika777@gmail.com³

ABSTRAK

Proses pengambilan data pada search engine yang masih memakai metode konvensional masih menyusahkan user dalam mencari dokumen yang diperlukan. Hasil dari pencarian itu masih berisi dokumen yang tidak relevant yang masih harus diseleksi secara manual oleh user. Clustering merupakan salah satu teknik dalam pengkategorian dokumen. Ide dasarnya merupakan dengan mengelompokkan dokumen-dokumen kedalam grup-grup atau cluster berdasarkan kemiripan (similarity) antar dokumen, sehingga dokumen yang berhubungan dengan suatu topik tertentu ditempatkan pada cluster yang sama. Dengan algoritma Fuzzy C-Means maka hasil pencarian hadits shahih akan di cluster atau dikelompokkan berdasarkan kemiripan kata antar dokumen. Berdasarkan pengujian yang telah dilakukan sistem dapat mengelompokkan hasil pencarian hadits shahih berdasarkan kemiripan kata. Selain itu, berdasarkan pengujian sistem dan pengujian manual menghasilkan bahwa algoritma fuzzy c-means memiliki tingkat akurasi sebesar 82,4% yang bekerja pada sistem pencarian hadits shahih.

Kata Kunci : *Fuzzy C-Means, Sistem Pencarian, Clustering, Search Engine*

ABSTRACT

The process of taking data on search engines that still use conventional methods still makes it difficult for users to find the documents needed. The results of that search still contain irrelevant documents that still need to be selected manually by the user. Clustering is one of the techniques in categorizing documents. The basic idea is to group documents into groups or clusters based on similarity (similarity) between documents, so documents related to a particular topic are placed in the same cluster. With Fuzzy C-Means algorithm, the results of the authentic hadith search will be clustered or grouped based on the similarity of words between documents. Based on testing that has been done the system can classify the results of the authentic hadith search based on word similarity. In addition, based on system testing and manual testing resulted that the fuzzy c-means algorithm has an accuracy rate of 82.4% that works on the authentic hadith search system.

Key Word : *Fuzzy C-Means, Search Systems, Clustering, Search Engines*

1. PENDAHULUAN

Proses pengambilan data pada search engine yang masih memakai metode konvensional masih menyusahkan user dalam mencari dokumen yang diperlukan. Hasil dari pencarian itu masih berisi dokumen yang tidak relevant yang masih harus diseleksi secara manual oleh user. Metode konvensional hanya mengurutkan hasil pencarian berdasarkan ranking. Sehingga diharapkan dengan adanya clustering, akan memudahkan user dalam melakukan pencarian, karena clustering akan mengelompokkan hasil. Juga sebagai sarana media pembelajaran dan referensi untuk belajar tentang Islam dari hadits shahih selain dari Al-Qur'an. Keanekaragaman dan pesatnya pertambahan jumlah dokumen dapat berpengaruh besar pada saat pencarian suatu dokumen. Suatu teknik pencarian dokumen yang ada pada saat ini merupakan hasil pencarian berurutan berdasarkan kecocokan kata atau kecocokan dokumen (*document ranking*). Pada referensi [1] menyatakan bahwa, terkadang hasil yang ditampilkan dengan pencarian berdasarkan *document ranking* terdiri dari beberapa halaman dan pengguna biasanya hanya melihat hasil pencarian maksimal pada 10 halaman pertama.

Pencarian dikelompokkan sesuai dengan kata yang diinputkan oleh pengguna. Dengan adanya pengelompokkan dokumen tersebut sangat menguntungkan bagi pengguna yang mencari lebih dari satu dokumen yang diinginkan selain itu pengguna tidak perlu membuka banyak halaman karena dokumen hasil pencarian dikelompokkan berdasarkan kategori yang mencerminkan isi dari suatu dokumen tersebut, sehingga pengguna dapat dimudahkan dalam pencarian dokumen yang diinginkan.

Tujuan dari dilakukannya penelitian ini yaitu untuk menerapkan metode FCM dalam pengelompokkan hasil pencarian hadits shahih serta untuk mengetahui kelompok manakah yang memiliki anggota terbaik pada pengelompokkan hasil pencarian hadits shahih.

Berdasarkan permasalahan yang ada berikut ini merupakan batasan masalah dari penelitian ini yaitu:

- a. Data yang digunakan dalam penelitian ini berupa data hadits shahih tentang rukun islam.
- b. Dalam penelitian ini menggunakan domain hadits shahih untuk melakukan klasterisasi data.
- c. Dalam Penelitian ini menggunakan hadits shahih yang di riwayatkan oleh An-Nasa'i.
- d. Klasterisasi data dengan menggunakan algoritma FCM dilakukan dengan mengelompokkan hasil pencarian hadits shahih.

Beberapa Penelitian sebelumnya, yaitu Menurut Rio Bayu A dan Lisa Yuli Kurniawati dalam penelitiannya mengusulkan sebuah metode baru untuk kategorisasi dokumen teks secara multi label dengan menggunakan Fuzzy C-Means dan K-Nearest Neighbors. Fuzzy C-Means melakukan pengelompokkan dokumen yang serupa terlebih dahulu sebelum proses pemberian label. Kemudian, penentuan label dokumen ditentukan oleh k-dokumen terdekat pada kelompok dokumen yang serupa. Uji coba dilakukan terhadap dokumen berita online sejumlah 175 dokumen yang terdiri atas tiga kategori label. Hasil uji coba menunjukkan bahwa metode yang diusulkan memberikan performa lebih baik dibanding metode lain. Hal ini ditunjukkan dengan nilai F1 sebesar 73,39% dan BEP sebesar 75,22%. [3].

Menurut Vidya Ayuningtias, dkk dalam penelitiannya melakukan pengelompokkan dokumen berbahasa Indonesia dengan algoritma K Means. Sebelum melakukan pengelompokkan diperlukan proses pre processing yaitu case folding, parsing, pembuangan stopword dan stemming. Proses ini diperlukan untuk mengurangi jumlah kata yang diproses. Nama kategori diberikan pada masing masing kluster. Penamaan kategori dihitung dari frekuensi kemunculan kata terbanyak dari setiap kluster. Analisis yang dilakukan pada jurnal ini merupakan membandingkan dua metode pembobotan dan korelasi jarak antar objek. Selain itu, analisis juga dilakukan untuk mengetahui kesesuaian nama kategori dengan dokumen dalam satu kluster dan menganalisis hasil kluster yang dihasilkan dengan melihat nilai convergence, precision dan recall. Hasilnya, Penamaan kategori dengan memilih maksimum frekuensi kata pada suatu kluster cukup relevan. Hal ini dibuktikan pada sub bab 5.2.6 yang menghasilkan

ketidakcocokan dokumen sebesar 28%. Selain itu, juga didukung oleh penilaian subjektif yang menyimpulkan bahwa 81% pengguna merasa terbantu dengan adanya penamaan kategori dalam pencarian dokumen. Keefektifitasan algoritma *k means* dalam pengelompokan dokumen pada jurnal ini mencapai nilai rata-rata *precision* 0,7464 dan nilai rata-rata *recall* 0,6750 yang diambil dari sepuluh kali percobaan. Sedangkan jika dilihat dari nilai maksimum dari sepuluh percobaan merupakan 0,8571 untuk *precision* dan 1 untuk *recall*. [2].

Data mining dapat diartikan sebagai kegiatan mengekstraksi atau menyaring suatu pengetahuan baru dari data yang berukuran atau berjumlah sangat besar yang dimana nantinya informasi ini sangatlah berguna untuk perkembangan suatu informasi.

Text mining dapat diartikan sebagai penemuan informasi yang baru dan tidak diketahui sebelumnya oleh komputer, dengan secara otomatis mengekstrak informasi dari sumber-sumber yang berbeda. Kunci dari proses ini merupakan menggabungkan informasi yang berhasil diekstraksi dari berbagai sumber (Hearst, 2003). Sedangkan menurut (Harlian Milkha, 2006) *text mining* memiliki definisi menambang data yang berupa teks dimana sumber data biasanya didapatkan dari dokumen, dan tujuannya merupakan mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen.

Klasterisasi data (*clustering*) merupakan sebuah proses untuk mengelompokkan data ke dalam beberapa klaster/kelompok sehingga data dalam satu klaster memiliki tingkat kemiripan yang maksimum dan data antar klaster memiliki kemiripan yang minimum (Tan, 2006). Implementasi klasterisasi ini dapat diterapkan pada berbagai bidang sebagai contoh dalam hal pemasaran, klasterisasi dapat digunakan sebagai metode untuk mengelompokkan pelanggan yang memiliki kesamaan dalam perilaku belanja. Penerapan klasterisasi dalam bidang ilmu biologi dapat dilakukan salah satunya dalam pengelompokan tumbuhan ataupun hewan berdasarkan fitur-fitur yang ada. Dalam bidang informatika sendiri, teknik klasterisasi sangat banyak digunakan antara lain pada data mining, information retrieval, pengolahan citra, dan lain sebagainya.

Fuzzy C-Means (FCM) merupakan salah satu algoritma Fuzzy Clustering. Fuzzy C-Means (FCM) merupakan suatu teknik pengklasteran data yang keberadaan setiap titik data dalam suatu cluster ditentukan oleh derajat keanggotaan. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981. Konsep dasar FCM yaitu menentukan pusat cluster, yang akan menandai lokasi rata-rata untuk setiap cluster. Dengan cara memperbaiki pusat cluster dan derajat keanggotaan setiap titik data secara berulang, maka akan dapat dilihat bahwa pusat cluster akan bergerak menuju lokasi yang tepat. Perulangan ini didasarkan pada minimasi fungsi objektif yang menggambarkan jarak dari titik data yang diberikan ke pusat cluster yang terbobot oleh derajat keanggotaan titik data tersebut. Output dari FCM bukan merupakan fuzzy inference system, namun merupakan deretan pusat cluster dan beberapa derajat keanggotaan untuk tiap-tiap titik data. Informasi ini dapat digunakan untuk membangun suatu fuzzy inference system.

Java merupakan bahasa pemrograman yang disusun oleh James Gosling yang dibantu oleh rekan-rekannya di suatu perusahaan perangkat lunak yang bernama Sun Microsystems, pada tahun 1991. Bahasa pemrograman ini mula-mula diinisialisasi dengan nama "Oak", namun pada tahun 1995 diganti namanya menjadi "Java". [12].

Metode Term Frequency-Inverse Document Frequency (TF-IDF) merupakan cara pemberian bobot hubungan suatu kata (term) terhadap dokumen. Untuk dokumen tunggal tiap kalimat dianggap sebagai dokumen.

Metode ini menggabungkan dua konsep untuk perhitungan bobot, yaitu Term frequency (TF) merupakan frekuensi kemunculan kata (t) pada kalimat (d). Document frequency (DF) merupakan banyaknya kalimat dimana suatu kata (t) muncul. Frekuensi kemunculan kata di dalam dokumen yang diberikan menunjukkan seberapa penting kata itu di dalam dokumen tersebut. Frekuensi dokumen yang mengandung kata tersebut menunjukkan seberapa umum kata tersebut. Bobot kata semakin besar jika sering muncul dalam suatu dokumen dan semakin kecil jika muncul dalam banyak dokumen (Robertson, 2005). Pada Metode ini pembobotan kata dalam sebuah dokumen dilakukan dengan mengalikan nilai TF dan IDF. Pembobotan diperoleh berdasarkan jumlah kemunculan term dalam kalimat (TF) dan jumlah kemunculan term pada

seluruh kalimat dalam dokumen (IDF). Bobot suatu istilah semakin besar jika istilah tersebut sering muncul dalam suatu dokumen dan semakin kecil jika istilah tersebut muncul dalam banyak dokumen (Grossman, 1998)

2. METODE PENELITIAN

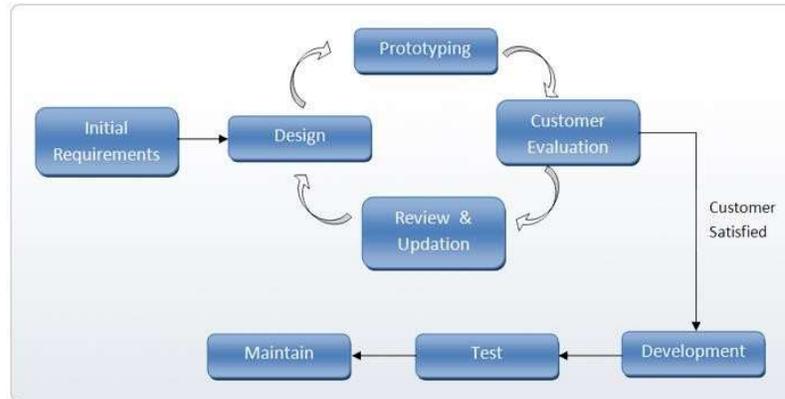
Dalam penelitian ini akan digunakan metode deskriptif dengan pendekatan kualitatif. Menurut Moleong, (2010 : 2) pendekatan kualitatif yaitu “data yang dikumpulkan cenderung bersifat naratif daripada angka-angka yang hasil analisisnya berupa uraian-uraian yang sangat deskriptif dan berdasarkan pada analisis data secara induktif”.

Selanjutnya Nasution (2008: 55-56) menyebutkan bahwa karakteristik penelitian kualitatif antara lain :“Pengambilan data yang dilakukan dalam suasana yang sewajarnya tanpa memanipulasi situasi yang ada dengan peneliti sebagai instrumen data, sumber bersifat purposive yakni diambil sesuai dengan fokus kajian, orang dapat memberikan informasi setulus mungkin; hasil penelitian berupa deskripsi, lebih mengutamakan proses dari produk; analisa data dilakukan secara terus menerus untuk mencari makna yang bersifat kontekstual atau sesuai dengan persepsi subjek yang diteliti; kesimpulan diraih melalui proses verifikasi.”

Dari pendapat tersebut bahwa sasaran yang dicapai dalam penelitian kualitatif diarahkan pada upaya menemukan teori-teori yang bersifat deskriptif. Moleong, (2010: 4-8) mengemukakan bahwa :“Prosesnya lebih diutamakan daripada hasil membatasi studinya dengan penentuan fokus dan menggunakan kriteria yang dipakai untuk kepentingan keabsahan data serta disepakati hasil penelitian oleh subyek penelitian dan peneliti”.

Pada tahap ini pengumpulan data dibagi menjadi dua tahapan yaitu yang pertama, tahap observasi merupakan pengumpulan suatu data dengan pendataan langsung dengan menganalisis dan menelaah data-data yang telah ada sebelumnya. Kemudian tahap yang kedua, studi literatur merupakan suatu teknik atau cara dalam pengumpulan suatu data dengan mempelajari literatur, internet, buku-buku, dokumentasi, modul, jurnal dan segala sesuatu yang dirasa perlu untuk memperkuat konsep dan teori yang dapat membantu menyelesaikan permasalahan sedang dibahas.

Metode Pengembangan Perangkat Lunak yang diterapkan dalam klasterisasi hasil pencarian menggunakan Prototype Model. Metode Prototype merupakan suatu paradigma baru dalam metode pengembangan perangkat lunak dimana metode ini tidak hanya sekedar evolusi dalam dunia pengembangan perangkat lunak, tetapi juga merevolusi metode pengembangan perangkat lunak yang lama yaitu sistem sekuensial yang biasa dikenal dengan nama SDLC atau waterfall development model. Dalam Model Prototype, prototype dari perangkat lunak yang dihasilkan kemudian dipresentasikan kepada user, dan user tersebut diberikan kesempatan untuk memberikan masukan sehingga perangkat lunak yang dihasilkan nantinya betul-betul sesuai dengan keinginan dan kebutuhan user. [4]. Metode ini menyediakan suatu konsep yang lengkap dari suatu sistem perangkat lunak, terdiri dari model kertas, model kerja dan program. Pengembang akan menganalisa kebutuhan pengguna, menganalisa sistem, melakukan studi kelayakan dan melakukan studi terhadap kebutuhan pemakai, meliputi model interface, teknik prosedural dan teknologi yang akan digunakan.

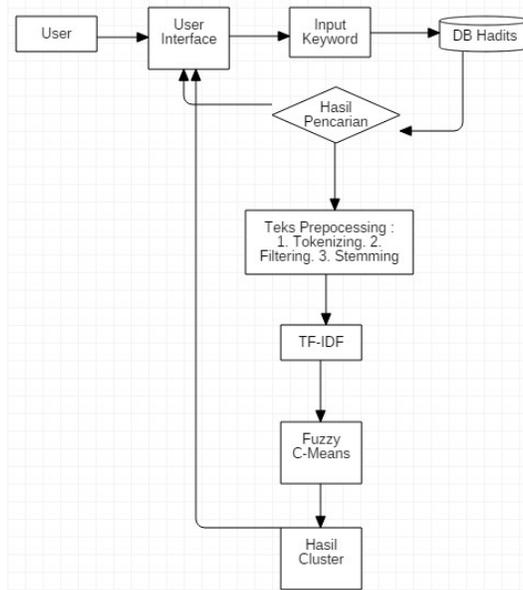


Gambar 1. Prototype Model

Berikut ini merupakan Tahapan-tahapan Proses Pengembangan Perangkat Lunak dengan Prototype yaitu :

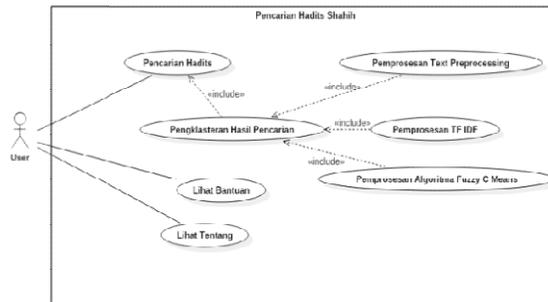
- a. Pengumpulan Kebutuhan
User dan pengembang bersama-sama mendefinisikan format seluruh perangkat lunak, mengidentifikasi semua kebutuhan, dan garis besar sistem yang akan dibuat.
- b. Membangun Prototyping
Membangun prototyping dengan membuat perancangan sementara yang berfokus pada penyajian kepada user (misalnya dengan membuat input dan format output).
- c. Evaluasi Prototyping
Evaluasi ini dilakukan oleh user, apakah prototyping yang sudah dibangun sudah sesuai dengan keinginan user atau belum. Jika sudah sesuai, maka langkah selanjutnya akan diambil. Namun jika tidak, prototyping direvisi dengan mengulang langkah-langkah sebelumnya.
- d. Pengkodean Sistem
Dalam tahap ini prototyping yang sudah di sepakati diterjemahkan ke dalam bahasa pemrograman yang sesuai.
- e. Pengujian Sistem
Setelah sistem sudah menjadi suatu perangkat lunak yang siap pakai, kemudian dilakukan proses Pengujian. Pengujian ini dilakukan dengan White Box, Black Box, Basis Path, pengujian arsitektur, dll.
- f. Evaluasi Sistem
User mengevaluasi apakah perangkat lunak yang sudah jadi sudah sesuai dengan yang diharapkan. Jika ya, maka proses akan dilanjutkan ke tahap selanjutnya, namun jika perangkat lunak yang sudah jadi tidak/belum sesuai dengan apa yang diharapkan, maka tahapan sebelumnya akan diulang.
- g. Penggunaan Sistem
Perangkat lunak yang telah diuji dan diterima user siap untuk digunakan.

Flowmap merupakan gambaran atau bentuk alur dari sistem yang dirancang. Untuk aplikasi pencarian dan pengclusteran hadis shahih ini , flowmap digambarkan sebagai berikut.



Gambar 2. Flowmap Arsitektur Pencarian Hadits Shahih

Use Case Diagram digunakan untuk memodelkan dan menyatakan unit fungsi yang disediakan oleh sistem kepada pengguna atau user. Use Case Diagram dari sistem ini dapat dilihat seperti pada gambar dibawah.



Gambar 3. Use Case Diagram Pencarian Hadits Shahih

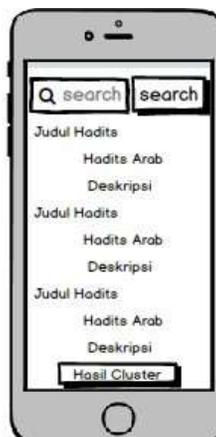
Rancangan antar muka (interface) aplikasi pengklasteran hasil pencarian hadits shahih :



Gambar 4. Interface Splash Screen



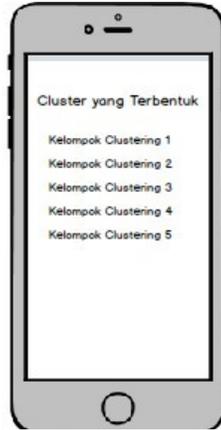
Gambar 5. *Interface* Menu Utama



Gambar 6. *Interface* Pencarian Hadits



Gambar 7. *Interface* Bantuan



Gambar 8. Interface Clustering



Gambar 9. Interface Hasil Clustering

Rancangan database pencarian hadits

Columns (6)

Column ID	Name	Type	Not Null	Default Value	Primary Key
0	id_hadits	INT(11)	1	null	0
1	judul_hadits	VARCHAR(255)	1	null	0
2	isi_arab	TEXT	0	null	0
3	sanad	TEXT	1	null	0
4	matan	TEXT	1	null	0
5	tipe	INTEGER	0	null	0

Gambar 10. Tabel Hadits

Columns (1)

Column ID	Name	Type	Not Null	Default Value	Primary Key
0	kata	TEXT	0	null	0

Gambar 11. Tabel Stopwords

Columns (2)

Column ID	Name	Type	Not Null	Default Value	Primary Key
0	name		0	null	0
1	seq		0	null	0

Gambar 12. Tabel sqlite_sequence

3. HASIL DAN PEMBAHASAN

Black box adalah pegujian sistem yang melakukan pengamatan hasil eksekusi dan memeriksa fungsionalitas perangkat lunak. Pengujian ini melakukan evaluasi berdasarkan apa yang dieksekusi melalui tampilan (*interface*) tanpa mengetahui proses yang terjadi didalam sebuah sistem perangkat lunak secara detail.

1. Pengujian Pencarian Hadits

Pengujian pencarian hadits merupakan pengujian fungsional untuk melakukan pencarian hadits yang diinginkan serta mendapatkan hasil pencarian berdasarkan *keyword* yang diinputkan oleh *user*. Pengujian pencarian hadits dapat dilihat pada tabel 1.

Tabel 1. Pengujian Pencarian Hadits

Kode	Skenario	Hasil		Keterangan
		Sukses	Gagal	
PS.01	<i>User</i> masuk ke <i>form</i> pencarian hadits.	√		Muncul form pencarian
PS.02	mengisi apa yang akan dicari pada <i>textbox</i> pencarian.	√		Tersedia <i>textbox</i> pencarian
PS.03	melihat hasil pencarian secara keseluruhan.	√		Menampilkan hasil pencarian keseluruhan

2. Pengujian Pengklasteran Hasil Pencarian

Pengujian pencarian hadits merupakan pengujian fungsional untuk melakukan pengklasteran hasil pencarian yang dilakukan oleh *user* berdasarkan *keyword*. Pengujian pengklasteran hasil pencarian dapat dilihat pada tabel 2.

Tabel 2. Pengujian Pengklasteran Hasil Pencarian

Kode	Skenario	Hasil		Keterangan
		Sukses	Gagal	
PS.04	memilih salah satu <i>cluster</i> atau kelompok dari hasil pencarian hadits.	√		Menampilkan nama kelompok <i>cluster</i>
PS.05	melihat hasil cluster dalam haditsnya.	√		Menampilkan hasil <i>cluster</i>

3. Pengujian Lihat Bantuan

Pengujian pencarian hadits merupakan pengujian fungsional untuk membuka *form* bantuan yang berisi informasi penggunaan perangkat lunak. Pengujian lihat bantuan dapat dilihat pada tabel 3

Tabel 3. Pengujian Lihat Bantuan

Kode	Skenario	Hasil		Keterangan
		Sukses	Gagal	
PS.06	User masuk ke form bantuan.	√		Muncul form bantuan
PS.07	melihat informasi form bantuan.	√		Menampilkan informasi tata cara pemakaian aplikasi

Berdasarkan hasil pengujian *black box* yang telah dilakukan dapat ditarik kesimpulan bahwa perangkat lunak yang dibangun sesuai dengan yang diharapkan dengan hasil sebuah perangkat lunak yang berjalan sesuai fungsionalitasnya.

Pengujian Algoritma *Fuzzy C-Means* dilakukan dengan membandingkan pengujian manual algoritma sebanyak 100 kali pengujian yang menggunakan 271 data hadits, menggunakan 10 buah sampel berupa *keyword* yang masing masing sampel diuji sebanyak 10 kali menggunakan proses *clustering* menggunakan Algoritma *Fuzzy C-Means*. Hasil pengujian algoritma *Fuzzy C-Means* dapat dilihat pada tabel 4.

Tabel 4. Pengujian Algoritma *Fuzzy C-Means*

No	Sampel (key)	Jumlah Pencarian	Hasil <i>Clustering</i> Sistem		Hasil <i>Clustering</i> Manual		Hasil
			Nama <i>Cluster</i>	Jumlah Anggota	Nama Cluster	Jumlah Anggota	
1	Shalat	60	Rakaat	27	1	25	92,5
			Hilang	9	2	12	75
			Matahar	24	3	23	95
2	zakat	25	Fitrah	3	1	3	100
			Ada	10	2	17	58
			Wajib	12	3	5	41
3	Puasa	64	Allah	36	1	36	100
			Buka	16	2	16	100
No	Sampel (key)	Jumlah Pencarian	Hasil <i>Clustering</i> Sistem		Hasil <i>Clustering</i> Manual		Hasil
			Nama <i>Cluster</i>	Jumlah Anggota	Nama Cluster	Jumlah Anggota	
			Lihat	12	3	12	
4	Haji	15	Orang	5	1	6	83

			Laku	4	2	3	75	
			Umrah	6	3	6	100	
5	Umrah	8	Haji	5	1	4	80	
			Orang	2	2	3	66	
			Hilang	1	3	1	100	
6	Buka	38	Puasa	17	1	16	94	
			Pintu	10	2	4	40	
			Lihat	11	3	18	61	
7	Matahari	18	Rasulul	5	1	4	80	
			Matahar	9	2	10	90	
			Shalat	4	3	4	100	
							Total	1730,5
							Hasil	

Dapat dilihat pada tabel 4 di kolom hasil didapat dari perhitungan dengan rumus dibawah ini :

$$\text{Hasil} = \frac{\text{Jumlah Anggota Manual}}{\text{Jumlah Anggota Sistem}} \times 100 \%$$

Rumus tersebut diimplementasikan perhitungan manual dan hasil pada sistem yang telah dilakukan dengan *keyword* yang sama dengan hasil *clustering* pada sistem. Hasil pengujian tersebut menghasilkan nilai akurasi berupa presentase kemudian dapat disimpulkan sebagai akurasi kinerja algoritma *fuzzy c-means*. Untuk mendapatkan nilai akurasi dapat menggunakan rumus diatas.

$$\text{Akurasi} = \frac{\text{Total Hasil}}{\text{Jumlah Percobaan}} \dots \dots \dots$$

$$\text{Akurasi} = \frac{1730,5}{21} = 82,4 \%$$

Akurasi merupakan rumus untuk menghitung akurasi yang dilakukan setelah melakukan pengujian dengan 7 sampel pengujian *clustering* dan menghasilkan 3 *cluster* didapatkanlah total akurasi sebesar 82,4%. Angka tersebut merupakan akurasi yang diberikan berdasarkan kinerja algoritma *fuzzy c-means* terhadap sistem.

4. KESIMPULAN

Dalam proses perancangan dan pembuatan perangkat lunak bantu pencarian hadits shahih dengan menggunakan Algoritma Fuzzy C-Means menghasilkan kesimpulan yang dapat penulis simpulkan sebagai berikut :

1. Algoritma Fuzzy C-Means yang digunakan dalam perangkat lunak bantu pengklasteran hasil pencarian hadits shahih memiliki tingkat akurasi sebesar 82,4%.
2. Kelompok cluster terbaik merupakan kelompok cluster yang memiliki anggota cluster yang memiliki jumlah anggota cluster yang sesuai antara perhitungan manual dengan hasil pada sistem.

5. SARAN

Mengingat berbagai keterbatasan yang dialami penulis diantaranya keterbatasan pemikiran dan waktu, maka penulis menyarankan untuk penelitian dimasa yang akan datang sebagai berikut :

1. Perlunya pengembangan dalam hal jumlah data yang diharapkan dapat ditambah dengan lebih banyak lagi.
2. Perlunya pengembangan dalam hal referensi kitab selain kitab sunan an'nasai.

DAFTAR PUSTAKA

- [1] B. Jansen dan U. Pooch, "A Review of Web Searching Studies and a Framework for Future Research," *Journal of the American Society for Information Science and Technology*, vol. 52, pp. 235-246, 2001.
- [2] V. Ayuningtyas, "Pengkategorian Hasil Pencarian Dokumen dengan Clustering," vol. 005.1, p. 8, 2008.
- [3] R. B. Afrianto dan L. Y. Kurniawati, "Kategorisasi Dokumen Teks Secara Multi Label Menggunakan Fuzzy C-Means Dan K-Nearest Neighbors Pada Artikel Berbahasa Indonesia," *Jurnal Ilmiah Teknologi Informasi*, vol. 1, 2013.
- [5] F. Zaenal dan S. Wiji, "'Pencarian dokumen Berbasis Web pada Drive Local dan Off-Line Web dengan Menggunakan Metode Suffix Tree Clustering'," *Jurnal Mahasiswa Fakultas Teknologi Informasi*, p. 7, 19 03 2015.
- [6] D. M. M. ., N. H. S. M. Bayu Wiradarma, "Implementasi Algoritma Fuzzy C-Means Untuk Pengelompokkan Tingkat Penyakit Anemia," *Jurnal Mahasiswa PTIIK*, vol. 7, 2016.
- [7] N. R. U. A. Entin Martiana, "Mesin Pencari Dokumen Dengan Pengklasteran Secara Otomatis," *ISSN: 1693-6930*, vol. 1, pp. 41-48, 2010.
- [8] I. Pramudiono, "Pengantar Data Mining: Menambang Permata Pengetahuan Di Gunung Data," *Kuliah Umum IlmuKomputer.Com*, 2003.
- [9] H. Februariyanti dan E. Zuliarso, "Klasifikasi Dokumen Berita Teks Bahasa Indonesia menggunakan Ontologi," *Jurnal Teknologi Informasi DINAMIK*, vol. 1, pp. 14-23, 2012.
- [10] R. Munir, dalam *Algoritma & Pemrograman dalam Pascal dan C*, Bandung, Informatika, 2011, p. 4.
- [11] M. Y. Soleh, "Implementasi Algoritma KMP dan Boyer-Moore dalam Aplikasi Search Engine Sederhana," *Makalah IF3051 Strategi Algoritma*, 2010/2011.
- [12] M. R. Arief, dalam *Pemrograman Web Dinamis Menggunakan Php dan Mysql*, Yogyakarta, ANDI, 2011.