

Analisis Pencocokan Nama dengan Nama Arab Terjemahan Bahasa Indonesia Menggunakan metode Levenshtein Distance

Wahyu Kurniawan¹, Moch. Arif Bijaksana², Bambang Ari Wahyudi³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹wahyukur93@gmail.com, ²arifbijaksana@telkomuniversity.ac.id, ³bambangari@telkomuniversity.ac.id

Abstrak

Dalam pencarian nama masih ada yang menuliskannya tidak sesuai dengan nama aslinya yang dicari sehingga nama yang ingin didapatkan itu tidak akan ketemu. Terkadang beberapa nama mempunyai kesamaan dalam pengucapannya, namun tulisannya berbeda-beda atau ejaannya pun bisa berbeda dari setiap orang yang menuliskannya. Karena nama yang ejaan atau tulisannya yang berbeda tersebut mungkin merupakan nama yang sama, sehingga perlu adanya penelitian nama yang berbeda tersebut memiliki kemiripan yang sama. Untuk mengetahui itu adalah nama yang sama, maka nama tersebut dicocokkan dengan metode *name matching* yang merupakan metode yang didalamnya terdapat algoritma untuk mencocokkan nama. Algoritma *levenshtein distance* merupakan salah satu metode yang digunakan dalam pencocokan nama yang dapat mengkonversi dua nama dengan tulisan atau ejaannya yang berbeda memiliki kemiripan yang sama. Hasil analisis yang didapat dari varian nama yang digunakan untuk pencocokkan nama menggunakan metode *levenshtein distance* mendapatkan nilai *precision*, *recall*, *f-measure* dan Akurasi yang berubah-ubah tergantung dari nilai kemiripannya dan pengaruh kesesuaian *query* dengan *gold standard* nya, sehingga nilai *precision*, *recall*, *f-measure* dan akurasi dapat mencapai 100%.

Kata kunci : *levenshtein distance*, *precision*, *recall*, *name matching*, akurasi, *f-measure*.

Abstract

In the search for names there are still those who write it not according to the original name they are looking for so that the name you want to get will not be found. Sometimes some names have similarities in pronunciation, but the writing is different or the spelling can be different from everyone who wrote it. Because the name whose spelling or writing is different may be the same name, so the need for research of different names has the same similarity. To find out that is the same name, the name is matched with the name matching method which is a method in which there is an algorithm to match names. The levenshtein distance algorithm is one of the methods used in matching names that can convert two names with different writing or spelling having the same similarity. The results of the analysis obtained from the name variants used for name matching using the levenshtein distance method get precision, recall, f-measure and accuracy that vary depending on the similarity value and the influence of the gold standard query compatibility, so the precision, recall, f-measure and accuracy can reach 100%.

Keywords: levenshtein distance, precision, recall, name matching, accuracy, f-measure.

1. Pendahuluan

Latar Belakang

Nama merupakan sebuah identitas yang dipunyai oleh seseorang. Setiap orang memiliki nama yang ejaannya berbeda-beda yang menjadikan ciri khas sendiri pada setiap orang. Dari nama kebanyakan memiliki kesamaan dalam pengucapannya, namun tulisannya berbeda-beda atau ejaannya pun bisa berbeda dari nama seperti "Muhammad" dengan "Mochammad" dan untuk mengetahui apakah nama yang ejaan dan tulisannya itu merupakan nama yang sama. Dalam tugas akhir ini akan membuat analisis untuk mengidentifikasi nama yang terkait dengan ejaan atau tulisannya yang berbeda untuk dicocokkan dengan nama arab dalam terjemahan bahasa indonesia sehingga diketahui bahwa nama yang ejaannya berbeda atau tulisannya berbeda bisa saja merupakan nama yang sama dan dapat dilihat saat mencari nama yang ada dalam hadits terutama mencari nama para perawi untuk mengetahui sejarah perawi serta periwayatannya[5].

Dalam tugas akhir ini dilakukan analisis untuk mencocokkan nama dari hasil nama yang diinputkan dengan nama yang berada di terjemahan hadits shahih bukhori. Dalam analisisnya menggunakan metode atau algoritma *levenshtein distance* yang merupakan algoritma untuk mencari kemiripan kata atau nama dengan mencari jaraknya

terlebih dahulu yaitu menghapus, menambah dan menyisipkan karakter yang tidak sama dalam pencocokan nama. Kemudian dicari kemiripannya dengan *levenshtein similarity*.

Dari penelitian yang lain terkait *name matching* menggunakan banyak metode untuk digunakan dalam menormalisasikan nama sebelum mencocokkan namanya, salah satunya yaitu dengan metode *levenshtein distance* untuk mencari jarak diantara kedua kata. Metode *levenshtein distance* merupakan metode yang dapat mencocokkan persamaan kata atau nama yang kemudian dapat menghitung kemiripan diantara kedua kata atau nama.

Topik dan Batasannya

Adapun perumusan masalah sesuai dengan latar belakang yang telah dibuat adalah untuk mengetahui bahwa nama yang dicari itu sama walaupun ejaannya berbeda dengan nama yang *dimatching* dan mencari nama yang mempunyai nilai kualitas mendekati jumlah data sebenarnya.

Memiliki batasan untuk mengerjakan tugas akhir ini mengenai topik tentang *name matching* yaitu mencocokkan nama dengan nama arab dari hadits shahih bukhori terjemahan bahasa indonesia, memakai metode *levenshtein distance* dalam mencocokkan nama arab, memfilter nama dengan huruf awal kapital, adanya daftar kata bukan nama untuk mengecek kata yang bukan nama, dataset yang dipakai adalah dari hadits shahih bukhori sejumlah 100 hadits, dapat mencari nama dengan 1 kata nama hingga 2 kata nama dengan kata awal Al, As, An, Asy, Ash, At, Ats, Ar, Az, Abu, Abdul, Ibnu dan kata kedua tidak dengan Al, As, An, Asy, At, Ats, Ar, Az, Abu.

Tujuan

Adapun tujuan yang telah dibuat sesuai dengan analisis yang telah dibuat yaitu untuk mengetahui bahwa nama yang dicari tersebut sama dengan nama yang *dimatching* menggunakan salah satu metode yaitu *Levenshtein distance* untuk menghitung seberapa mirip nama tersebut dan mengetahui nilai kualitas terhadap nama yang dicocokkan dengan menggunakan *precision, recall, f-measure* dan akurasi.

2. Studi Terkait

2.1 Hadits

Hadits adalah sesuatu hal tentang nabi yang berisi tentang semua yang dilakukan atau berkaitan oleh nabi dapat berupa segala hal tentang nabi atau bisa juga tentang kehidupan nabi sebelum atau sesudah menjadi nabi. Hadist dapat berupa sesuatu yang menjadi referensi yang kemudian diterapkan, dilakukan atau dituliskan kembali [1].

2.2 Name Matching

Name matching ini bukan hanya merupakan komponen di proses identitas *matching*. Dua data yang ditunjukkan dapat berbeda namun ditunjukkan bahwa nama itu sama yaitu antara "Barack" dan "Husein". Nama keluarga dapat juga menjadikan bahwa kedua orang tersebut merupakan orang yang sama contohnya dari keluarga yang sama yaitu "Obama". Dalam perbandingan nama belum tentu cukup untuk membuktikan identitas itu sendiri [6].

Name matching adalah salah satu unsur yang merupakan identitas *matching* yang berdampak langsung pada pergeseran yang terjadi ke arah *linguistic matching* yang menjadikan teks berfokus ke cara nama yang dicocokkan. Nama juga hanya memiliki satu identitas yang mungkin juga bisa dicocokkan dengan pencariannya [6].

2.3 Lavenshtein distance

Salah satu langkah paling populer di *name matching* yang jumlah minimum operasi untuk mengkonversi nama [6]. Operasi dasarnya yaitu :

- Penyisipan karakter (misalnya Yeltsin → Yexltsin)
- Penghapusan karakter (misalnya Yeltsin → Yeltsn)
- Penggantian dengan karakter yang lain (misalnya Yeltsin → Yqltsin)

Algoritma *levenshtein distance* memperbandingkan dua kata yang kemudian diukur berdasarkan berdasarkan kata masing-masing yang kemudian dihitung dengan menggunakan matrik yang dibangun dengan baris dan kolom dapat digambarkan dengan n dan m yang berisi kata per karakter masing-masing. Memakai fungsi yang mengembalikkan nilai 0 jika karakter yang dicocokkan sama dan sebaliknya memberikan nilai 1 jika karakter yang

dicocokkan tidak sama [4]. Jarak Edit dikenal adalah Jarak *levenshtein distance* (setelah matematika Rusia Vladimir Levenshtein) [6]. Dan perhitungan dasarnya yaitu :

$$\text{Levenshtein distance}(S1, S2) = 1.0 - \frac{\text{dist}(S1, S2)}{\max(|S1|, |S2|)} \quad (1)$$

Di mana $\text{dist}(S1, S2)$ adalah yang sebenarnya jarak *levenshtein distance* yang mengandung nilai 0 jika *string* yang satu jumlahnya positif jika keduanya tidak sama. Ukuran nilai di antara 0.0 dan 1.0 ketika ada nilai yang lebih besar akan semakin banyak yang didapat di antara dua *string* yang sama [2].

2.4 Matching Quality

Pencocokan itu memberikan sebuah kemiripan, namun harus memiliki kualitas dengan banyaknya matching sesuai entitas yang sebenarnya. Adanya kategori yang akan sesuai dengan kecocokannya dari dua namanya yaitu TP (*true positive*) adalah pasangan nama yang benar-benar cocok dan merupakan kecocokan keduanya, FP (*false positive*) adalah pasangan nama yang awalnya memang cocok dan merupakan kecocokan palsu, TN (*true negative*) adalah pasangan nama yang tidak cocok dan memang tidak cocok, dan FN (*false negative*) adalah pasangan nama yang awalnya cocok dan kebenarannya tidak cocok. Dari kategori tersebut akan dihitung *precision* dan *recall* [2] yaitu

$$\text{Precision} : \frac{TP}{TP + FP}, \text{Recall} : \frac{TP}{TP + FN} \quad (2)$$

Dihitung juga *F-measure* yang didapat dari menggabungkan *precision and recall* dengan rata-rata harmoniknya yaitu

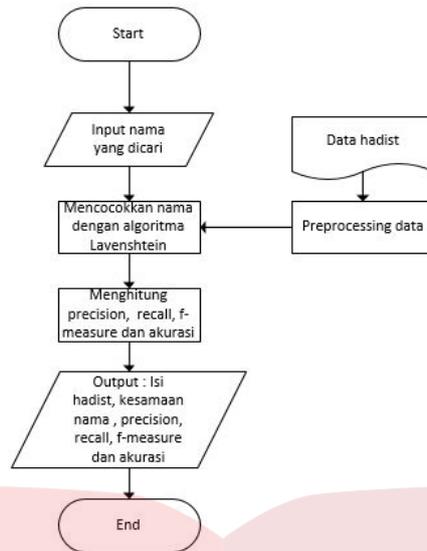
$$F\text{-measure} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

Dengan memanfaatkan TP (*true positive*), FP (*false positive*), FN (*false negative*), TN (*true negative*) yang ada pada *precision and recall*, nilai perhitungan akurasi bisa didapatkan. Adapun cara menghitung akurasi [3] yaitu

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

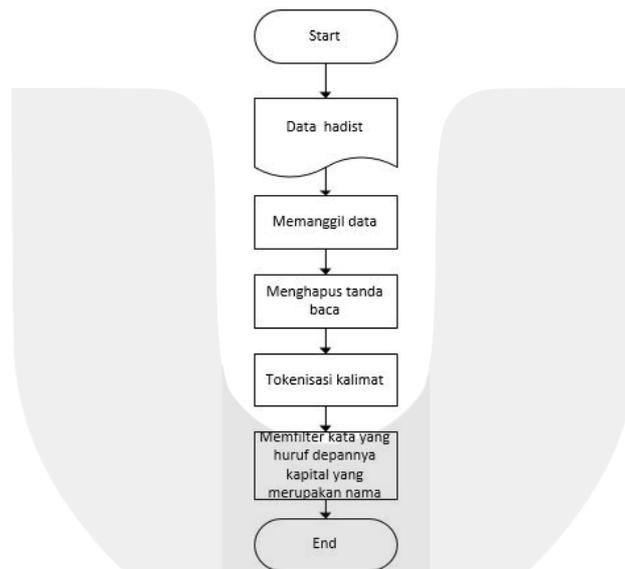
3. Sistem yang Dibangun

Sistem yang dibangun ini adalah sistem untuk mencocokkan nama dengan nama arab terjemahan bahasa indonesia yang terdapat dari dataset hadits shahih bukhori menggunakan metode *levenshtein distance*. Dengan metode *levenshtein distance* ini dapat melakukan pencocokan nama yang kemudian menghasilkan jarak pencocokkannya dan kemiripannya. Sebelum mencocokkan nama tersebut data hadits harus diolah terlebih dahulu untuk menjadi sebuah sebuah nama yang kemudian dapat dicocokkan. Melakukan *preprocessing* data yaitu didalamnya terdapat proses memanggil data hadits, menghapus semua tanda baca, mentokenisasi dan memfilter berdasarkan huruf awal kapital untuk mencari kata yang memungkinkan mengandung nama yang itu semua merupakan proses untuk menjadikan data tersebut siap untuk dicocokkan. Sistem menginputkan nama dengan satu kata atau dua kata, kemudian yang mengandung tanda "-" diganti menjadi " " untuk 2 kata dan data hadits yang sudah diproses, kemudian dicocokkan dengan metode *levenshtein distance* untuk mencari kemiripan dari pencocokan kedua namanya. Setelah itu dihitunglah *precision*, *recall*, *f-measure* dan akurasi dari hasil kecocokan inputan dengan *gold standard* kemudian menghasilkan *output* berupa isi hadits, kemiripan nama dan hasil dari *precision*, *recall*, *f-measure* dan akurasi yang didapat. Gambaran sistem yang dibangun sebagai berikut dan untuk contohnya dapat dilihat di lampiran.



Gambar 1. Gambaran sistem yang dibuat

Proses *preprocessing* data lebih detail untuk sistem yang dibangun



Gambar 2. Gambaran *preprocessing* data

1. Memanggil data yaitu proses memanggil data dari hadits shahih bukhori yang berjumlah 100 yang dalam bentuk file .txt
2. Menghapus tanda baca yaitu proses menghapus tanda baca yang terkandung dalam hadits shahih bukhori yang berjumlah 100
3. Tokenisasi kalimat yaitu proses memecah-mecah kalimat di hadits menjadi kata
4. Memfilter kata yang hurufnya kapital yaitu proses untuk mengumpulkan kata-kata yang awalnya kapital dan mengecek kata yang bukan nama dari daftar kata bukan nama agar tidak muncul, sehingga dapat memunculkan kata yang merupakan nama.

4. Evaluasi

4.1 Skenario Pengujian 1

4.1.1 Hasil Pengujian

Pengujian dilakukan untuk mencari *precision*, *recall*, *f-measure* dan akurasi dari hasil inputan *query* yang didapat melalui kuesioner yang dibagikan ke responden. Pengujian dilakukan dengan mencocokkan inputan responden (*query*) dengan nama yang ada di hadits shahih bukhori berjumlah 100 dan yang kemudian oleh sistem dicari kemiripannya dan dicocokkan dengan *gold standard* yang sudah diverifikasi oleh DR. Agus Suyadi Raharusun Lc, M.Ag selaku dosen prodi Ilmu Hadis, Fakultas Ushuluddin, UIN Sunan Gunung Djati Bandung untuk mencari nilai *precision*, *recall*, *f-measure* dan akurasi. Di bawah ini adalah hasil pengujian dengan kemiripan untuk 1 kata nama yang ditampilkan yaitu $\geq 70\%$, $\geq 75\%$ dan $\geq 80\%$ yang terlihat di tabel 1 dan tabel 2. Untuk lebih lengkapnya dapat dilihat di lampiran yaitu lampiran 4 untuk pengujian dari kemiripan 70% sampai 80%.

Tabel 1. Hasil pengujian 1 kata nama (Kemiripan, TP, FP, FN, TN)

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Batas kemiripan	Kemiripan	TP	FP	FN	TN
1	عَائِشَةَ	Aisyah	Aisyah	13	$\geq 70\%$	100.00%	13	0	0	1466
		'Aisyah				85.71%	13	0	0	1466
		Aissyah				85.71%	13	0	0	1466
	بِلَالٍ	Bilal	Bilal	4		100.00%	4	2	0	1473
		Bilaal				83.33%	4	0	0	1475
		Biilal				83.33%	4	0	0	1475
	عَبَّاسٍ	Abbas	Abbas	25		100.00%	25	0	0	1454
		A'bbaas				71.43%	25	0	0	1454
		Abas				80.00%	25	23	0	1431
2	عَائِشَةَ	Aisyah	Aisyah	13	$\geq 75\%$	100.00%	13	0	0	1466
		'Aisyah				85.71%	13	0	0	1466
		Aissyah				85.71%	13	0	0	1466
	بِلَالٍ	Bilal	Bilal	4		100.00%	4	2	0	1473
		Bilaal				83.33%	4	0	0	1475
		Biilal				83.33%	4	0	0	1475
	عَبَّاسٍ	Abbas	Abbas	25		100.00%	25	0	0	1454
		A'bbaas				0.00%	0	0	0	1479
		Abas				80.00%	25	23	0	1431
3	عَائِشَةَ	Aisyah	Aisyah	13	$\geq 80\%$	100.00%	13	0	0	1466
		'Aisyah				85.71%	13	0	0	1466
		Aissyah				85.71%	13	0	0	1466
	بِلَالٍ	Bilal	Bilal	4		100.00%	4	2	0	1473
		Bilaal				83.33%	4	0	0	1475
		Biilal				83.33%	4	0	0	1475
	عَبَّاسٍ	Abbas	Abbas	25		100.00%	25	0	0	1454
		A'bbaas				0.00%	0	0	0	1479
		Abas				80.00%	25	0	0	1454

Tabel 2. Hasil pengujian 1 kata nama (*Precision, Recall, F-measure, Akurasi*)

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Batas kemiripan	Precision	Recall	F-measure	Akurasi
1	عَائِشَة	Aisyah	Aisyah	13	$\geq 70\%$	100.00%	100.00%	100.00%	100.00%
		'Aisyah				100.00%	100.00%	100.00%	100.00%
		Aissyah				100.00%	100.00%	100.00%	100.00%
	بِلَال	Bilal	Bilal	4		66.67%	100.00%	80.00%	99.86%
		Bilaal				100.00%	100.00%	100.00%	100.00%
		Biilal				100.00%	100.00%	100.00%	100.00%
	عَبَّاس	Abbas	Abbas	25		100.00%	100.00%	100.00%	100.00%
		A'bbaas				100.00%	100.00%	100.00%	100.00%
		Abas				52.08%	100.00%	68.49%	98.44%
2	عَائِشَة	Aisyah	Aisyah	13	$\geq 75\%$	100.00%	100.00%	100.00%	100.00%
		'Aisyah				100.00%	100.00%	100.00%	100.00%
		Aissyah				100.00%	100.00%	100.00%	100.00%
	بِلَال	Bilal	Bilal	4		66.67%	100.00%	80.00%	99.86%
		Bilaal				100.00%	100.00%	100.00%	100.00%
		Biilal				100.00%	100.00%	100.00%	100.00%
	عَبَّاس	Abbas	Abbas	25		100.00%	100.00%	100.00%	100.00%
		A'bbaas				0.00%	0.00%	0.00%	100.00%
		Abas				52.08%	100.00%	68.49%	98.44%
3	عَائِشَة	Aisyah	Aisyah	13	$\geq 80\%$	100.00%	100.00%	100.00%	100.00%
		'Aisyah				100.00%	100.00%	100.00%	100.00%
		Aissyah				100.00%	100.00%	100.00%	100.00%
	بِلَال	Bilal	Bilal	4		66.67%	100.00%	80.00%	99.86%
		Bilaal				100.00%	100.00%	100.00%	100.00%
		Biilal				100.00%	100.00%	100.00%	100.00%
	عَبَّاس	Abbas	Abbas	25		100.00%	100.00%	100.00%	100.00%
		A'bbaas				0.00%	0.00%	0.00%	100.00%
		Abas				100.00%	100.00%	100.00%	100.00%

4.1.2 Analisis Hasil Pengujian

Bahwa dari hasil pengujian yang dilakukan menggunakan metode *levenshtein distance* dengan kemiripan dari 70% sampai 80% yang diambil untuk perwakilan pengujian yaitu kemiripan $\geq 70\%$, $\geq 75\%$ dan $\geq 80\%$ yang diperbandingkan dengan *gold standard* yang sudah diverifikasi. Dilihat pengujian menunjukkan, bahwa *precision, recall, f-measure* dan Akurasinya 100% itu pengaruh dari *query* sesuai dengan jumlah di *gold standard* dan dapat terlihat di TP, FP, FN dan TN. Untuk nilai *precision, recall, f-measure* dan Akurasinya kurang dari 100%, karena *query* tidak sesuai dengan jumlah di *gold standard* dan dapat terlihat di TP, FP, FN, TN. Ketika *precision, recall, f-measure* bernilai 0% atau bernilai 100%, karena nilai kemiripannya kurang dari perwakilan pengujian kemiripan yaitu $\geq 70\%$, $\geq 75\%$ dan $\geq 80\%$.

4.2 Skenario pengujian 2

4.2.1 Hasil Pengujian

Pengujian dilakukan untuk mencari *precision, recall, f-measure* dan akurasi dari hasil inputan *query* yang didapat melalui kuesioner yang dibagikan ke responden. Pengujian dilakukan dengan mencocokkan inputan responden dengan nama yang ada di hadits shahih bukhori berjumlah 100 hadits dan yang kemudian oleh sistem dicari kemiripannya dengan metode *levenshtein distance* dan dicocokkan dengan *gold standard* yang sudah diverifikasi oleh DR. Agus Suyadi Raharusun Lc, M.Ag selaku dosen prodi Ilmu Hadis, Fakultas Ushuluddin, UIN Sunan Gunung Djati Bandung untuk mencari nilai *precision, recall, f-measure* dan akurasi. Di bawah ini adalah hasil pengujian dengan kemiripan untuk 2 kata nama yang ditampilkan yaitu $\geq 70\%$, $\geq 75\%$ dan $\geq 80\%$ yang terlihat di tabel 3 dan tabel 4. Untuk lebih lengkapnya dapat dilihat di lampiran yaitu lampiran 5 untuk pengujian dari kemiripan 70% sampai 80%.

Tabel 3. Hasil pengujian 2 kata nama (Kemiripan, TP, FP, FN, TN)

No	Bahasa Arabnya	Query	Goldstandard	Batas kemiripan	Kemiripan	TP	FP	FN	TN
1	الزُّهْرِيُّ	Az Zuhri	Az Zuhri	$\geq 70\%$	100.00%	17	0	0	305
		Az-Zuhriy			88.89%	17	0	0	305
	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah		71.43%	18	0	0	304
		Abi Hurairah			91.67%	18	0	0	304
	إِبْنُ شِهَابٍ	Ibnu Syihab	Ibnu Syihab		100.00%	14	2	0	306
		Ibnu Syihaabin			78.57%	14	0	0	308
2	الزُّهْرِيُّ	Az Zuhri	Az Zuhri	$\geq 75\%$	100.00%	17	0	0	305
		Az-Zuhriy			88.89%	17	0	0	305
	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah		0.00%	0	0	0	322
		Abi Hurairah			91.67%	18	0	0	304
	إِبْنُ شِهَابٍ	Ibnu Syihab	Ibnu Syihab		100.00%	14	0	0	308
		Ibnu Syihaabin			78.57%	14	0	0	308
3	الزُّهْرِيُّ	Az Zuhri	Az Zuhri	$\geq 80\%$	100.00%	17	0	0	305
		Az-Zuhriy			88.89%	17	0	0	305
	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah		0.00%	0	0	0	322
		Abi Hurairah			91.67%	18	0	0	304
	إِبْنُ شِهَابٍ	Ibnu Syihab	Ibnu Syihab		100.00%	14	0	0	308
		Ibnu Syihaabin			0.00%	0	0	0	322

Tabel 4. Hasil pengujian 2 kata nama (*Precision, Recall, F-measure, Akurasi*)

No	Bahasa Arabnya	Query	Goldstandard	Batas kemiripan	Precision	Recall	F-measure	Akurasi
1	الزُّهْرِيُّ	Az Zuhri	Az Zuhri	$\geq 70\%$	100.00%	100.00%	100.00%	100.00%
		Az-Zuhriy			100.00%	100.00%	100.00%	100.00%
	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah		100.00%	100.00%	100.00%	100.00%
		Abi Hurairah			100.00%	100.00%	100.00%	100.00%
	إِبْنُ شِهَابٍ	Ibnu Syihab	Ibnu Syihab		87.50%	100.00%	93.33%	99.38%
		Ibnu Syihaabin			100.00%	100.00%	100.00%	100.00%
2	الزُّهْرِيُّ	Az Zuhri	Az Zuhri	$\geq 75\%$	100.00%	100.00%	100.00%	100.00%
		Az-Zuhriy			100.00%	100.00%	100.00%	100.00%
	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah		0.00%	0.00%	0.00%	100.00%
		Abi Hurairah			100.00%	100.00%	100.00%	100.00%
	إِبْنُ شِهَابٍ	Ibnu Syihab	Ibnu Syihab		100.00%	100.00%	100.00%	100.00%
		Ibnu Syihaabin			100.00%	100.00%	100.00%	100.00%
3	الزُّهْرِيُّ	Az Zuhri	Az Zuhri	$\geq 80\%$	100.00%	100.00%	100.00%	100.00%
		Az-Zuhriy			100.00%	100.00%	100.00%	100.00%
	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah		0.00%	0.00%	0.00%	100.00%
		Abi Hurairah			100.00%	100.00%	100.00%	100.00%
	إِبْنُ شِهَابٍ	Ibnu Syihab	Ibnu Syihab		100.00%	100.00%	100.00%	100.00%
		Ibnu Syihaabin			0.00%	0.00%	0.00%	100.00%

4.2.2 Analisis Hasil Pengujian

Hasil dari pengujian yang dilakukan dengan kemiripan dari yang diambil untuk perwakilan pengujian yaitu kemiripan $\geq 70\%$, $\geq 75\%$ dan $\geq 80\%$ dengan metode *levenshtein distance* yang diperbandingkan dengan *gold standard* yang sudah diverifikasi oleh ahlinya. Dilihat dari pengujian, bahwa nilai *precision*, *recall*, *f-measure* dan akurasi 100% itu pengaruh dari *query* yang dicocokkan sesuai dengan jumlah di *gold standard* dan dapat terlihat juga di TP, FP, FN dan TN. Untuk nilai *precision*, *recall*, *f-measure* dan akurasi kurang dari 100% itu pengaruh dari *query* yang dicocokkan dengan jumlah yang di *gold standard* tidak sesuai dan dapat terlihat juga di TP, FP, FN dan TN. Sedangkan nilai *precision*, *recall*, *f-measure* 0%, karena nilai kemiripannya tidak sesuai dengan batasan kemiripannya, sehingga nilai kemiripannya 0%.

5. Kesimpulan

Kesimpulan yang dapat diambil dalam penelitian yang telah dilakukan yaitu kinerja sistem untuk mencocokkan nama dengan menggunakan metode *levenshtein distance* mendapatkan nilai *precision*, *recall*, *f-measure* dan Akurasi yang berubah-ubah tergantung dari nilai kemiripannya dan pengaruh kesesuaian *query* dengan *gold standard*nya, sehingga nilai *precision*, *recall*, *f-measure* dan akurasi dapat mencapai 100%.

Saran yang dapat diambil yaitu bisa dikembangkan lagi untuk dataset selain sejumlah 100 hadits shahih bukhori dan membuat pencocokan nama tidak hanya terbatas untuk nama yang terdiri 2 kata dengan kata awal Al, As, An, Asy, Ash, At, Ats, Ar, Az, Abu, Abdul, Ibnu dan kata kedua tidak dengan Al, As, An, Asy, At, Ats, Ar, Az, Abu.

Daftar Pustaka

- [1] S. M. Al-Qaththan. *Pengantar Studi Ilmu Hadits*. Pustaka Al Kautsar, 2012.
- [2] M. ALIFIKRI. *INDONESIAN NAME MATCHING USING MACHINE LEARNING SUPERVISED APPROACH*. Telkom University, 2017.
- [3] S. P. Benny, S. Vasavi, and P. Anupriya. Hadoop framework for entity resolution within high velocity streams. *Procedia Computer Science*, 85:550–557, 2016.
- [4] A. T. Freeman, S. L. Condon, and C. M. Ackerman. Cross linguistic name matching in english and arabic: a one to many mapping extension of the levenshtein edit distance algorithm. In *Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*, pages 471–478. Association for Computational Linguistics, 2006.
- [5] A. Herdi. *Memahami Ilmu Hadis*. Tafakur, 2014.
- [6] B. Lisbach and V. Meyer. *Linguistic identity matching*. Springer, 2013.

Lampiran

Lampiran 1

Contoh ingin mencocokkan nama, namun data harus diproses seperti dibawah ini

1. Menghapus tanda baca pada kalimat dari hadist shahih bukhori
 - (a) Contoh menggunakan kalimat : telah menceritakan kepada kami Abdullah bin Yusuf berkata,
 - (b) Maka dihapuslah tanda baca tersebut menjadi : telah menceritakan kepada kami Abdullah bin Yusuf berkata
2. Memisah-misah kalimat menjadi kata seperti pada gambar 3



Gambar 3. Gambaran Proses Memisah-misah Kalimat

3. Memfilter kata yang mengandung huruf awal Kapital seperti pada gambar 4 dan mengecek kata yang bukan nama dari daftar kata bukan nama yang sudah dibuat, contohnya : Telah, Apa, Ada, dll, agar kata tersebut tidak muncul.



Gambar 4. Gambaran Proses Memfilter Kata Kapital

4. Mencocokkan nama dengan metode *Lavenshtein distance* yaitu dengan inputan Abdulah dan Ibnu Syihaabin
 - (a) Mencari jarak antara kedua nama
 - i. Pertama Abdulah dengan Abdullah 1 proses yaitu
 - A. Menyisipkan karakter l di Abdulah menjadi Abdullah
 - ii. Kedua dengan nama yang lain yaitu Ibnu Syihaabin dengan Ibnu Syihab 3 proses yaitu
 - A. Menghapus karakter a di Ibnu Syihaabin menjadi Ibnu Syihabin
 - B. Menghapus karakter i di Ibnu Syihabin menjadi Ibnu Syihabn
 - C. Menghapus karakter n di Ibnu Syihabin menjadi Ibnu Syihab
 - iii. Mencari jarak kedua nama yaitu Abdulah dengan Abdullah menggunakan matriks
 - A. Pertama yaitu masukkan nama Abdulah ke dalam tabel matriks secara vertikal dan nama Abdullah secara horisontal. Pencocokkan ini untuk merubah nama Abdul menjadi Abdullah yang dapat dilihat pada gambar 5.
 - B. Kedua yaitu lakukan penambahan huruf sampai tidak menjadi kosong (null) dari mempunyai huruf a, ab sampai menjadi abdullah dan menambah nilai +1 untuk penambahan huruf seperti pada baris ke 2 yang dapat dilihat pada gambar 6.
 - C. Ketiga yaitu lakukan pengurangan huruf sampai tidak menjadi kosong (null) dari huruf a, ab sampai menjadi abdullah menjadi tidak ada huruf (null) dan menambah nilai +1 untuk pengurangan huruf seperti pada kolom ke 2 yang dapat dilihat pada gambar 7.

	Null	a	b	d	u	l	l	a	h
Null									
a									
b									
d									
u									
l									
a									
h									

Gambar 5. Langkah pertama

	Null	a	b	d	u	l	l	a	h
Null	0	1	2	3	4	5	6	7	8
a									
b									
d									
u									
l									
a									
h									

Gambar 6. Langkah kedua

	Null	a	b	d	u	l	l	a	h
Null	0	1	2	3	4	5	6	7	8
a	1								
b	2								
d	3								
u	4								
l	5								
a	6								
h	7								

Gambar 7. Langkah ketiga

- D. Keempat yaitu lakukan pencocokkan antar huruf dengan huruf yang ada pada baris sama huruf yang ada pada kolom. Jika ada perbedaan antar huruf, maka cari nilai terkecil dari sel di kirinya, sel di kiri atasnya sama sel di atasnya, lalu ditambah 1 seperti di tabel matriks yang membandingkan huruf a sama b yaitu ambil nilai terkecil dari sel di kirinya ($sel(1,2) = 1$), sel di kiri atasnya ($sel(2,2) = 0$) dan sel di atasnya ($sel(1,3) = 2$), lalu ditambah 1, kemudian hasilnya adalah di $sel(2,3) = 1$. Jika ada persamaan antar huruf, maka ambil nilai dari sel di kiri atasnya seperti pada tabel matriks membandingkan huruf l dengan l di $sel(6,6)$ nilainya menjadi 0 yang dapat dilihat pada gambar 8.

	Null	a	b	d	u	l	l	a	h
Null	0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	4	5	6	7
b	2	1	0	1	2	3	4	5	6
d	3	2	1	0	1	2	3	4	5
u	4	3	2	1	0	1	2	3	4
l	5	4	3	2	1	0	1	2	3
a	6								
h	7								

Gambar 8. Langkah keempat

- E. Kelima yaitu lakukan hal sebelumnya sampai tabel matriks menjadi penuh dan didapatkan *distance* dari membandingkan nama abduallah dengan abduallah adalah 1 yang terlihat di pojok kanan bawah yang dapat dilihat pada gambar 9.

	Null	a	b	d	u	l	l	a	h
Null	0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	4	5	6	7
b	2	1	0	1	2	3	4	5	6
d	3	2	1	0	1	2	3	4	5
u	4	3	2	1	0	1	2	3	4
l	5	4	3	2	1	0	1	2	3
a	6	5	4	3	2	1	1	1	2
h	7	6	5	4	3	2	2	2	1

Gambar 9. Langkah kelima

- iv. Mencari jarak kedua nama yaitu Ibnu Syihaabin dengan Ibnu Syihab menggunakan matriks dengan pencarian yang sama seperti mencari Abdulah dengan Abdullah dan hasilnya dapat dilihat di gambar 10.

	null	i	b	n	u	s	y	i	h	a	b
null											
i											
b											
n											
u											
s											
y											
i											
h											
a											
a											
b											
i											
n											

	null	i	b	n	u	s	y	i	h	a	b
null	0	1	2	3	4	5	6	7	8	9	10
i	1	0	1	2	3	4	5	6	7	8	9
b	2	1	0	1	2	3	4	5	6	7	8
n	3	2	1	0	1	2	3	4	5	6	7
u	4	3	2	1	0	1	2	3	4	5	6
s	5	4	3	2	1	0	1	2	3	4	5
y	6	5	4	3	2	1	0	1	2	3	4
i	7	6	5	4	3	2	1	0	1	2	3
h	8	7	6	5	4	3	2	1	0	1	2
a	9	8	7	6	5	4	3	2	1	0	1
a	10	9	8	7	6	5	4	3	2	1	0
b	11	10	9	8	7	6	5	4	3	2	1
b	12	11	10	9	8	7	6	5	4	3	2
i	13	12	11	10	9	8	7	6	5	4	3
n	14	13	12	11	10	9	8	7	6	5	4

Gambar 10. Tabel Pencarian Jarak dengan Matriks 2

(b) Menghitung levenshtein similarity

- i. Pertama Abdulah dengan Abdullah

$$Levenshtein(S1, S2) = 1.0 - \frac{distance(S1, S2)}{maxLength(S1, S2)} = 1.0 - \frac{1}{8} = 0.875 * 100\% = 87.5\% \quad (5)$$

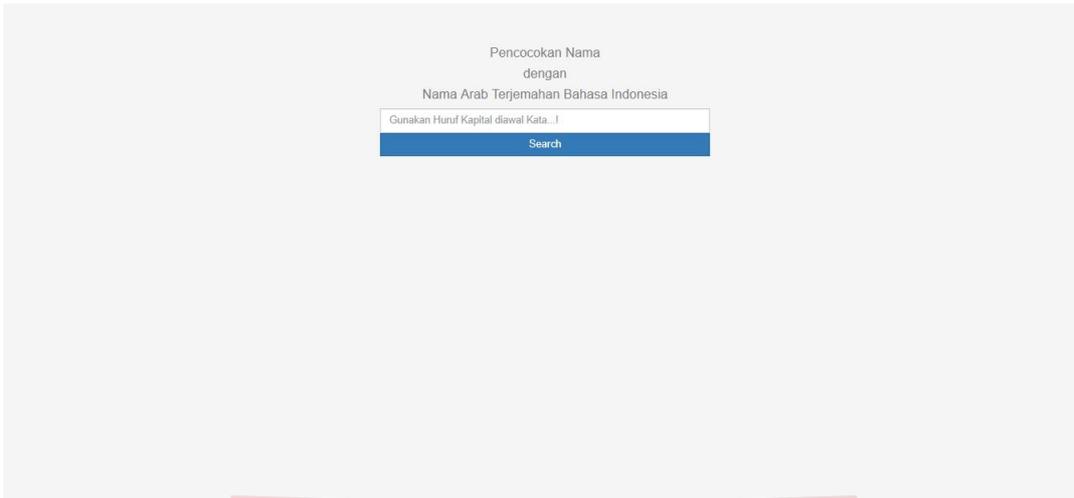
- ii. Kedua Ibnu Syihaabin dengan Ibnu Syihab

$$Levenshtein(S1, S2) = 1.0 - \frac{distance(S1, S2)}{maxLength(S1, S2)} = 1.0 - \frac{3}{14} = 0.786 * 100\% = 78.6\% \quad (6)$$

- 5. Mencari precision and recall dari banyaknya kemunculan kecocokan yang kemiripannya >= 0.7 dan nama di gold standard
- 6. Mencari f-measure dari 2 dikali dengan hasil dari pembagian yang dilakukan yaitu (precision*recall) dan (precision+recall)
- 7. Mencari akurasi diambil dari TP ditambah dengan TN kemudian dibagi dengan hasil dari penjumlahan TP, TN, FP dan FN

Lampiran 2

Berikut adalah gambar-gambar screenshot aplikasi yang digunakan untuk mencocokkan nama yaitu di gambar 11 pertama menginputkan nama yang ingin dicari, kemudian klik search dan hasilnya menampilkan isi hadist yang kemunculannya berdasarkan keluarnya nama yang dicocokkan dengan batasan similarity, nilai similarity, jumlah gold standard, precision, recall, f-measure dan akurasi di gambar 12. Untuk gambar 13 adalah tampilan pencarian aplikasi yang bukan nama dan gambar 14 adalah tampilan pencarian aplikasi 2 kata nama.



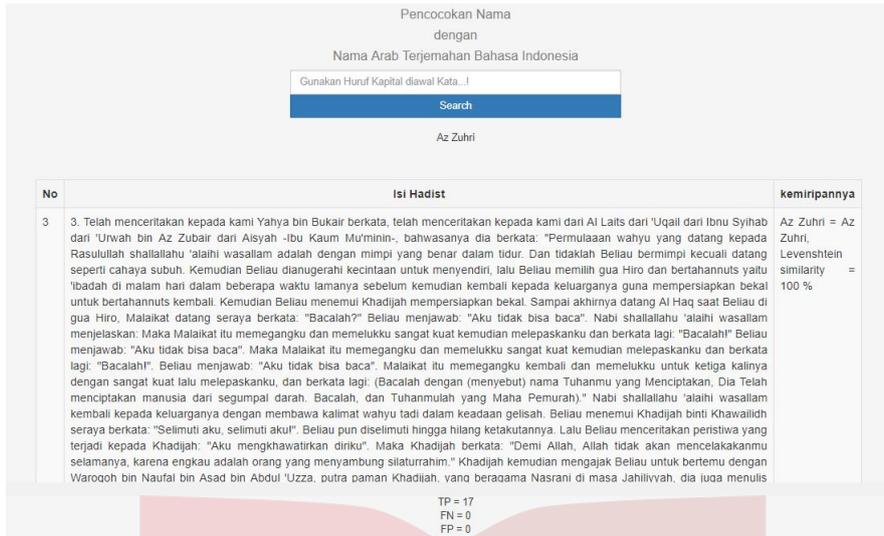
Gambar 11. Tampilan Awal Aplikasi



Gambar 12. Tampilan aplikasi dijalankan



Gambar 13. Tampilan pencarian aplikasi yang bukan nama



Gambar 14. Tampilan pencarian aplikasi untuk 2 kata nama

Lampiran 3

Berikut adalah *gold standard* yang ditampilkan terpisah yaitu pada link : <https://drive.google.com/file/d/1MIFgI6ifvT7M4TqGxzUCrQMjRjOsc2iE/view?usp=sharing> dan gambar 15 adalah surat pernyataan tentang *gold standard* yang sudah diverifikasi.



Gambar 15. Surat Verifikasi

Lampiran 4

Berikut hasil pengujian untuk pencocokkan 1 kata nama lebih detail dari kemiripan 70% sampai 80% yang ditunjukkan pada gambar 16 sampai gambar 26.

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi	
1	أبي	Abu	Abu	145	145	100.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%	
2		Abuu			145	75.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%	100.00%
3		Abū			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%	100.00%
4	عائشة	Aisyah	Aisyah	13	13	100.00%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%	
5		'Aisyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%	
6		Aissyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%	
7	قَتَادَة	Qotadah	Qotadah	9	9	100.00%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%	
8		Qotaadah			9	87.50%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%	
9		Qotaadata			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%	
10	مُحَمَّد	Muhammad	Muhammad	52	52	100.00%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%	
11		Muhammadu			52	88.89%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%	
12		Muhamad			52	87.50%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%	
13	عُمَر	Umar	Umar	21	21	100.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%	
14		'Umar			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%	
15		Umaro			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%	
16	بِلَال	Bilal	Bilal	4	6	100.00%	4	2	0	1473	66.67%	100.00%	80.00%	99.86%	
17		BilaaI			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%	
18		Biial			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%	
19	عَبَّاس	Abbas	Abbas	25	25	100.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%	
20		A'bbas			25	71.43%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%	
21		Abas			48	80.00%	25	23	0	1431	52.08%	100.00%	68.49%	98.44%	
22	عَبْدُ اللَّهِ	Abdullah	Abdullah	82	95	100.00%	82	13	0	1396	86.32%	100.00%	92.66%	99.12%	
23		A'bdullah			82	88.89%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%	
24		Abdulloh			82	87.50%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%	
25	سُؤْبَاه	Syub'ah	Syubah	22	24	85.71%	22	2	0	1455	91.67%	100.00%	95.65%	99.86%	
26		Syukbah			24	85.71%	22	2	0	1455	91.67%	100.00%	95.65%	99.86%	
27		Syubah			24	100.00%	22	2	0	1455	91.67%	100.00%	95.65%	99.86%	
28	عَبْدُ الرَّحْمَنِ	Abdurrahman	Abdurrahman	11	12	100.00%	11	1	0	1467	91.67%	100.00%	95.65%	99.93%	
29		Abdurrokhman			11	83.33%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%	
30		'Abdurrochmani			11	71.43%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%	
Rata-rata nilai						83.03%					89.06%	93.33%	90.79%	99.90%	

Gambar 16. Pengujian pencocokkan 1 kata nama dengan batasan kemiripan 70%

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	أبي	Abu	Abu	145	145	100.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
2		Abuu			145	75.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
3		Abū			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
4	عائشة	Aisyah	Aisyah	13	13	100.00%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
5		'Aisyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
6		Aissyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
7	قَتَادَة	Qotadah	Qotadah	9	9	100.00%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
8		Qotaadah			9	87.50%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
9		Qotaadata			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
10	مُحَمَّد	Muhammad	Muhammad	52	52	100.00%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
11		Muhammadu			52	88.89%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
12		Muhamad			52	87.50%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
13	عُمَر	Umar	Umar	21	21	100.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
14		'Umar			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
15		Umaro			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
16	بِلَال	Bilal	Bilal	4	6	100.00%	4	2	0	1473	66.67%	100.00%	80.00%	99.86%
17		BilaaI			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
18		Biial			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
19	عَبَّاس	Abbas	Abbas	25	25	100.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
20		A'bbas			25	71.43%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
21		Abas			48	80.00%	25	23	0	1431	52.08%	100.00%	68.49%	98.44%
22	عَبْدُ اللَّهِ	Abdullah	Abdullah	82	83	100.00%	82	1	0	1396	98.80%	100.00%	99.39%	99.93%
23		A'bdullah			82	88.89%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
24		Abdulloh			82	87.50%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
25	سُؤْبَاه	Syub'ah	Syubah	22	24	85.71%	22	2	0	1455	91.67%	100.00%	95.65%	99.86%
26		Syukbah			24	85.71%	22	2	0	1455	91.67%	100.00%	95.65%	99.86%
27		Syubah			24	100.00%	22	2	0	1455	91.67%	100.00%	95.65%	99.86%
28	عَبْدُ الرَّحْمَنِ	Abdurrahman	Abdurrahman	11	12	100.00%	11	1	0	1467	91.67%	100.00%	95.65%	99.93%
29		Abdurrokhman			11	83.33%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
30		'Abdurrochmani			11	71.43%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
Rata-rata nilai						83.03%					89.47%	93.33%	91.02%	99.93%

Gambar 17. Pengujian pencocokkan 1 kata nama dengan batasan kemiripan 71%

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	أبي	Abu	Abu	145	145	100.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
2		Abuu			145	75.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
3		Abū			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
4	عائشة	Aisyah	Aisyah	13	13	100.00%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
5		'Aisyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
6		Aissyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
7	قَتَادَة	Qotadah	Qotadah	9	9	100.00%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
8		Qotaadah			9	87.50%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
9		Qotaadata			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
10	مُحَمَّدٌ	Muhammad	Muhammad	52	52	100.00%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
11		Muhammadu			52	88.89%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
12		Muhamad			52	87.50%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
13	عُمَرُ	Umar	Umar	21	21	100.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
14		'Umar			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
15		Umaro			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
16	بِلَالٌ	Bilal	Bilal	4	6	100.00%	4	2	0	1473	66.67%	100.00%	80.00%	99.86%
17		Bilaa			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
18		Biilal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
19	عَبَّاسٌ	Abbas	Abbas	25	25	100.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
20		A'bbaas			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
21		Abas			48	80.00%	25	23	0	1431	52.08%	100.00%	68.49%	98.44%
22	عَبْدُ اللَّهِ	Abdullah	Abdullah	82	83	100.00%	82	1	0	1396	98.80%	100.00%	99.39%	99.93%
23		A'bdullah			82	88.89%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
24		Abdulloh			82	87.50%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
25	شُعْبَةَ	Syubah	Syubah	22	22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
26		Syukbah			22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
27		Syubah			22	100.00%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
28	عَبْدُ الرَّحْمَنِ	Abdurrahman	Abdurrahman	11	12	100.00%	11	1	0	1467	91.67%	100.00%	95.65%	99.93%
29		Abdurrohman			11	83.33%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
30		'Abdurrochmani			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						78.27%					83.64%	86.67%	84.78%	99.94%

Gambar 18. Pengujian pencocokkan 1 kata nama dengan batasan kemiripan 72%

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	أبي	Abu	Abu	145	145	100.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
2		Abuu			145	75.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
3		Abū			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
4	عائشة	Aisyah	Aisyah	13	13	100.00%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
5		'Aisyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
6		Aissyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
7	قَتَادَة	Qotadah	Qotadah	9	9	100.00%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
8		Qotaadah			9	87.50%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
9		Qotaadata			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
10	مُحَمَّدٌ	Muhammad	Muhammad	52	52	100.00%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
11		Muhammadu			52	88.89%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
12		Muhamad			52	87.50%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
13	عُمَرُ	Umar	Umar	21	21	100.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
14		'Umar			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
15		Umaro			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
16	بِلَالٌ	Bilal	Bilal	4	6	100.00%	4	2	0	1473	66.67%	100.00%	80.00%	99.86%
17		Bilaa			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
18		Biilal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
19	عَبَّاسٌ	Abbas	Abbas	25	25	100.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
20		A'bbaas			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
21		Abas			48	80.00%	25	23	0	1431	52.08%	100.00%	68.49%	98.44%
22	عَبْدُ اللَّهِ	Abdullah	Abdullah	82	83	100.00%	82	1	0	1396	98.80%	100.00%	99.39%	99.93%
23		A'bdullah			82	88.89%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
24		Abdulloh			82	87.50%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
25	شُعْبَةَ	Syubah	Syubah	22	22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
26		Syukbah			22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
27		Syubah			22	100.00%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
28	عَبْدُ الرَّحْمَنِ	Abdurrahman	Abdurrahman	11	11	100.00%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
29		Abdurrohman			11	83.33%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
30		'Abdurrochmani			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						78.27%					83.92%	86.67%	84.93%	99.94%

Gambar 19. Pengujian pencocokkan 1 kata nama dengan batasan kemiripan 73%

Lampiran 5

Berikut hasil pengujian untuk pencocokkan 2 kata nama lebih detail dari kemiripan 70% sampai 80% yang ditunjukkan pada gambar 27 sampai gambar 37.

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi	
1	أبي	Abu	Abu	145	145	100.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%	
2		Abuu			145	75.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%	100.00%
3		Abū			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%	100.00%
4	عائشة	Aisyah	Aisyah	13	13	100.00%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%	
5		'Aisyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%	
6		Aissyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%	
7	قَتَادَة	Qotadah	Qotadah	9	9	100.00%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%	
8		Qotaadah			9	87.50%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%	
9		Qotaadata			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%	
10	مُحَمَّدٌ	Muhammad	Muhammad	52	52	100.00%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%	
11		Muhammadu			52	88.89%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%	
12		Muhamad			52	87.50%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%	
13	عُمَر	Umar	Umar	21	21	100.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%	
14		'Umar			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%	
15		Umara			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%	
16	بِلَال	Bilal	Bilal	4	6	100.00%	4	2	0	1473	66.67%	100.00%	80.00%	99.86%	
17		Bilaaal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%	
18		Biilal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%	
19	عَبَّاس	Abbas	Abbas	25	25	100.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%	
20		A'bbaas			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%	
21		Abas			48	80.00%	25	23	0	1431	52.08%	100.00%	68.49%	98.44%	
22	عَبْدُ اللَّهِ	Abdullah	Abdullah	82	83	100.00%	82	1	0	1396	98.80%	100.00%	99.39%	99.93%	
23		'Abdullah			82	88.89%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%	
24		Abdulloh			82	87.50%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%	
25	شُعْبَة	Syubah	Syubah	22	22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%	
26		Syukbah			22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%	
27		Syubah			22	100.00%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%	
28	عَبْدُ الرَّحْمَنِ	Abdurrahman	Abdurrahman	11	11	100.00%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%	
29		Abdurrokhman			11	83.33%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%	
30		'Abdurrochmani			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%	
Rata-rata nilai						78.27%					83.92%	86.67%	84.93%	99.94%	

Gambar 20. Pengujian pencocokkan 1 kata nama dengan batasan kemiripan 74%

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	أبي	Abu	Abu	145	145	100.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
2		Abuu			145	75.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
3		Abū			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
4	عائشة	Aisyah	Aisyah	13	13	100.00%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
5		'Aisyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
6		Aissyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
7	قَتَادَة	Qotadah	Qotadah	9	9	100.00%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
8		Qotaadah			9	87.50%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
9		Qotaadata			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
10	مُحَمَّدٌ	Muhammad	Muhammad	52	52	100.00%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
11		Muhammadu			52	88.89%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
12		Muhamad			52	87.50%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
13	عُمَر	Umar	Umar	21	21	100.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
14		'Umar			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
15		Umara			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
16	بِلَال	Bilal	Bilal	4	6	100.00%	4	2	0	1473	66.67%	100.00%	80.00%	99.86%
17		Bilaaal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
18		Biilal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
19	عَبَّاس	Abbas	Abbas	25	25	100.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
20		A'bbaas			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
21		Abas			48	80.00%	25	23	0	1431	52.08%	100.00%	68.49%	98.44%
22	عَبْدُ اللَّهِ	Abdullah	Abdullah	82	83	100.00%	82	1	0	1396	98.80%	100.00%	99.39%	99.93%
23		'Abdullah			82	88.89%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
24		Abdulloh			82	87.50%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
25	شُعْبَة	Syubah	Syubah	22	22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
26		Syukbah			22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
27		Syubah			22	100.00%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
28	عَبْدُ الرَّحْمَنِ	Abdurrahman	Abdurrahman	11	11	100.00%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
29		Abdurrokhman			11	83.33%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
30		'Abdurrochmani			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						78.27%					83.92%	86.67%	84.93%	99.94%

Gambar 21. Pengujian pencocokkan 1 kata nama dengan batasan kemiripan 75%

Lampiran 6

Gambar 38 dan gambar 39 dibawah ini adalah hasil dari kuesioner yang digunakan untuk pengujian 1 kata nama yang nantinya dibuat kasus uji

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	أبي	Abu	Abu	145	145	100.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
2		Abuu			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
3		Abū			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
4	عائشة	Aisyah	Aisyah	13	13	100.00%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
5		'Aisyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
6		Aissyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
7	قَتَادَة	Qotadah	Qotadah	9	9	100.00%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
8		Qotaadah			9	87.50%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
9		Qotaadata			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
10	مُحَمَّدٌ	Muhammad	Muhammad	52	52	100.00%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
11		Muhammadu			52	88.89%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
12		Muhamad			52	87.50%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
13	عُمَر	Umar	Umar	21	21	100.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
14		'Umar			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
15		Umaro			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
16	بِلَال	Bilal	Bilal	4	6	100.00%	4	2	0	1473	66.67%	100.00%	80.00%	99.86%
17		Bilaaal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
18		Biilal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
19	عَبَّاس	Abbas	Abbas	25	25	100.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
20		A'bbaas			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
21		Abas			25	80.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
22	عَبْدُ اللَّهِ	Abdullah	Abdullah	82	83	100.00%	82	1	0	1396	98.80%	100.00%	99.39%	99.93%
23		A'bdullah			82	88.89%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
24		Abdulloh			82	87.50%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
25	شُعْبَةَ	Syubah	Syubah	22	22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
26		Syukbah			22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
27		Syubah			22	100.00%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
28	عَبْدُ الرَّحْمَنِ	Abdurrahman	Abdurrahman	11	11	100.00%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
29		Abdurrohkman			11	83.33%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
30		'Abdurrochmani			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						75.77%					82.18%	83.33%	82.65%	99.99%

Gambar 22. Pengujian pencocokkan 1 kata nama dengan batasan kemiripan 76%

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	أبي	Abu	Abu	145	145	100.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
2		Abuu			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
3		Abū			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
4	عائشة	Aisyah	Aisyah	13	13	100.00%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
5		'Aisyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
6		Aissyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
7	قَتَادَة	Qotadah	Qotadah	9	9	100.00%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
8		Qotaadah			9	87.50%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
9		Qotaadata			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
10	مُحَمَّدٌ	Muhammad	Muhammad	52	52	100.00%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
11		Muhammadu			52	88.89%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
12		Muhamad			52	87.50%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
13	عُمَر	Umar	Umar	21	21	100.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
14		'Umar			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
15		Umaro			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
16	بِلَال	Bilal	Bilal	4	6	100.00%	4	2	0	1473	66.67%	100.00%	80.00%	99.86%
17		Bilaaal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
18		Biilal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
19	عَبَّاس	Abbas	Abbas	25	25	100.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
20		A'bbaas			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
21		Abas			25	80.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
22	عَبْدُ اللَّهِ	Abdullah	Abdullah	82	83	100.00%	82	1	0	1396	98.80%	100.00%	99.39%	99.93%
23		A'bdullah			82	88.89%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
24		Abdulloh			82	87.50%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
25	شُعْبَةَ	Syubah	Syubah	22	22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
26		Syukbah			22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
27		Syubah			22	100.00%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
28	عَبْدُ الرَّحْمَنِ	Abdurrahman	Abdurrahman	11	11	100.00%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
29		Abdurrohkman			11	83.33%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
30		'Abdurrochmani			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						75.77%					82.18%	83.33%	82.65%	99.99%

Gambar 23. Pengujian pencocokkan 1 kata nama dengan batasan kemiripan 77%

Lampiran 7

Gambar 40 dibawah ini adalah hasil dari kuesioner yang digunakan untuk pengujian 2 kata nama yang nantinya dibuat kasus uji

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	أبي	Abu	Abu	145	145	100.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
2		Abuu			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
3		Abū			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
4	عائشة	Aisyah	Aisyah	13	13	100.00%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
5		'Aisyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
6		Aissyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
7	قَتَادَة	Qotadah	Qotadah	9	9	100.00%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
8		Qotaadah			9	87.50%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
9		Qotaadata			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
10	مُحَمَّدٌ	Muhammad	Muhammad	52	52	100.00%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
11		Muhammadu			52	88.89%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
12		Muhamad			52	87.50%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
13	عُمَر	Umar	Umar	21	21	100.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
14		'Umar			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
15		UmARO			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
16	بِلَال	Bilal	Bilal	4	6	100.00%	4	2	0	1473	66.67%	100.00%	80.00%	99.86%
17		Bilaal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
18		Biilal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
19	عَبَّاس	Abbas	Abbas	25	25	100.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
20		A'bbaas			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
21		Abas			25	80.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
22	عَبْدُ اللَّهِ	Abdullah	Abdullah	82	82	100.00%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
23		A'bdullah			82	88.89%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
24		Abdulloh			82	87.50%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
25	شُعْبَة	Syu'bah	Syubah	22	22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
26		Syukbah			22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
27		Syubah			22	100.00%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
28	عَبْدُ الرَّحْمَنِ	Abdurrahman	Abdurrahman	11	11	100.00%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
29		Abdurrokhman			11	83.33%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
30		'Abdurrochmani			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						75.77%					82.22%	83.33%	82.67%	100.00%

Gambar 24. Pengujian pencocokkan 1 kata nama dengan batasan kemiripan 78%

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	أبي	Abu	Abu	145	145	100.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
2		Abuu			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
3		Abū			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
4	عائشة	Aisyah	Aisyah	13	13	100.00%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
5		'Aisyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
6		Aissyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
7	قَتَادَة	Qotadah	Qotadah	9	9	100.00%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
8		Qotaadah			9	87.50%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
9		Qotaadata			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
10	مُحَمَّدٌ	Muhammad	Muhammad	52	52	100.00%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
11		Muhammadu			52	88.89%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
12		Muhamad			52	87.50%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
13	عُمَر	Umar	Umar	21	21	100.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
14		'Umar			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
15		UmARO			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
16	بِلَال	Bilal	Bilal	4	6	100.00%	4	2	0	1473	66.67%	100.00%	80.00%	99.86%
17		Bilaal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
18		Biilal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
19	عَبَّاس	Abbas	Abbas	25	25	100.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
20		A'bbaas			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
21		Abas			25	80.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
22	عَبْدُ اللَّهِ	Abdullah	Abdullah	82	82	100.00%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
23		A'bdullah			82	88.89%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
24		Abdulloh			82	87.50%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
25	شُعْبَة	Syu'bah	Syubah	22	22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
26		Syukbah			22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
27		Syubah			22	100.00%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
28	عَبْدُ الرَّحْمَنِ	Abdurrahman	Abdurrahman	11	11	100.00%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
29		Abdurrokhman			11	83.33%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
30		'Abdurrochmani			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						75.77%					82.22%	83.33%	82.67%	100.00%

Gambar 25. Pengujian pencocokkan 1 kata nama dengan batasan kemiripan 79%

Lampiran 8

Untuk daftar kata yang bukan nama dapat dibuka dilink ini : <https://drive.google.com/file/d/1EohjzLdJhVlUVl0gEDY2b0o1RXeJytT2/view?usp=sharing> , karena tidak mencukupi untuk ditampilkan.

No	Bahasa Arabnya	Query	Goldstandar	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	أبي	Abu	Abu	145	145	100.00%	145	0	0	1334	100.00%	100.00%	100.00%	100.00%
2		Abuu			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
3		Abū			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
4	عائشة	Aisyah	Aisyah	13	13	100.00%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
5		'Aisyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
6		Aissyah			13	85.71%	13	0	0	1466	100.00%	100.00%	100.00%	100.00%
7	قَتَادَة	Qotadah	Qotadah	9	9	100.00%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
8		Qotaadah			9	87.50%	9	0	0	1470	100.00%	100.00%	100.00%	100.00%
9		Qotaadata			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
10	مُحَمَّدٌ	Muhammad	Muhammad	52	52	100.00%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
11		Muhammadu			52	88.89%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
12		Muhamad			52	87.50%	52	0	0	1427	100.00%	100.00%	100.00%	100.00%
13	عُمَر	Umar	Umar	21	21	100.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
14		'Umar			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
15		Umaro			21	80.00%	21	0	0	1458	100.00%	100.00%	100.00%	100.00%
16	بِلَال	Bilal	Bilal	4	6	100.00%	4	2	0	1473	66.67%	100.00%	80.00%	99.86%
17		Bilaaal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
18		Biilal			4	83.33%	4	0	0	1475	100.00%	100.00%	100.00%	100.00%
19	عَبَّاس	Abbas	Abbas	25	25	100.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
20		A'bbas			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
21		Abas			25	80.00%	25	0	0	1454	100.00%	100.00%	100.00%	100.00%
22	عَبْدُ اللَّهِ	Abdullah	Abdullah	82	82	100.00%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
23		'abdullah			82	88.89%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
24		Abdulloh			82	87.50%	82	0	0	1397	100.00%	100.00%	100.00%	100.00%
25	شُعْبَة	Syubah	Syubah	22	22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
26		Syukbah			22	85.71%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
27		Syubah			22	100.00%	22	0	0	1457	100.00%	100.00%	100.00%	100.00%
28	عَبْدُ الرَّحْمَنِ	Abdurrahman	Abdurrahman	11	11	100.00%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
29		Abdurrokhman			11	83.33%	11	0	0	1468	100.00%	100.00%	100.00%	100.00%
30		'Abdurrochmani			0	0.00%	0	0	0	1479	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						75.77%					82.22%	83.33%	82.67%	100.00%

Gambar 26. Pengujian pencocokkan 1 kata nama dengan batasan kemiripan 80%

No	Bahasa Arabnya	Query	Goldstandar	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	الرُّهْرِي	Az Zuhri	Az Zuhri	17	17	100.00%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
2		Az-Zuhriy			17	88.89%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
3	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah	18	18	71.43%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
4		Abi Hurairah			18	91.67%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
5	أَبِي سُفْيَانَ	Abii Sufyaan	Abu Sufyan	10	10	75.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
6		Abi Sufyan			10	90.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
7	ابْنُ سَهَابٍ	Ibnu Syihab	Ibnu Syihab	14	16	100.00%	14	2	0	306	87.50%	100.00%	93.33%	99.38%
8		Ibnu Syihaabin			14	78.57%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
9	ابْنُ عَبَّاسٍ	Ibnu Abbas	Ibnu Abbas	19	19	100.00%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
10		Ibnu 'Abbaasin			19	71.43%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
Rata-rata nilai						86.70%					98.75%	100.00%	99.33%	99.94%

Gambar 27. Pengujian pencocokkan 2 kata nama dengan batasan kemiripan 70%

No	Bahasa Arabnya	Query	Goldstandar	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	الرُّهْرِي	Az Zuhri	Az Zuhri	17	17	100.00%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
2		Az-Zuhriy			17	88.89%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
3	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah	18	18	71.43%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
4		Abi Hurairah			18	91.67%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
5	أَبِي سُفْيَانَ	Abii Sufyaan	Abu Sufyan	10	10	75.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
6		Abi Sufyan			10	90.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
7	ابْنُ سَهَابٍ	Ibnu Syihab	Ibnu Syihab	14	16	100.00%	14	2	0	306	87.50%	100.00%	93.33%	99.38%
8		Ibnu Syihaabin			14	78.57%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
9	ابْنُ عَبَّاسٍ	Ibnu Abbas	Ibnu Abbas	19	19	100.00%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
10		Ibnu 'Abbaasin			19	71.43%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
Rata-rata nilai						86.70%					98.75%	100.00%	99.33%	99.94%

Gambar 28. Pengujian pencocokkan 2 kata nama dengan batasan kemiripan 71%

No	Bahasa Arabnya	Query	Goldstandar	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	الرُّهْرِي	Az Zuhri	Az Zuhri	17	17	100.00%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
2		Az-Zuhriy			17	88.89%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
3	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah	18	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
4		Abi Hurairah			18	91.67%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
5	أَبِي سُفْيَانَ	Abii Sufyaan	Abu Sufyan	10	10	75.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
6		Abi Sufyan			10	90.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
7	ابْنُ سَهَابٍ	Ibnu Syihab	Ibnu Syihab	14	16	100.00%	14	2	0	306	87.50%	100.00%	93.33%	99.38%
8		Ibnu Syihaabin			14	78.57%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
9	ابْنُ عَبَّاسٍ	Ibnu Abbas	Ibnu Abbas	19	19	100.00%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
10		Ibnu 'Abbaasin			0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						72.41%					78.75%	80.00%	79.33%	99.94%

Gambar 29. Pengujian pencocokkan 1 kata nama dengan batasan kemiripan 72%

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	الرُّهْرِيّ	Az Zuhri	Az Zuhri	17	17	100.00%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
2		Az-Zuhriy			17	88.89%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
3	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah	18	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
4		Abi Hurairah			18	91.67%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
5	أَبِي سُفْيَانَ	Abii Sufyaan	Abu Sufyan	10	10	75.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
6		Abi Sufyan			10	90.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
7	ابْنُ سَيِّهَابٍ	Ibnu Syihab	Ibnu Syihab	14	14	100.00%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
8		Ibnu Syihaabin			14	78.57%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
9	ابْنُ عَبَّاسٍ	Ibnu Abbas	Ibnu Abbas	19	19	100.00%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
10		Ibnu 'Abbaasin			0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						72.41%					80.00%	80.00%	80.00%	100.00%

Gambar 30. Pengujian pencocokkan 2 kata nama dengan batasan kemiripan 73%

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	الرُّهْرِيّ	Az Zuhri	Az Zuhri	17	17	100.00%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
2		Az-Zuhriy			17	88.89%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
3	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah	18	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
4		Abi Hurairah			18	91.67%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
5	أَبِي سُفْيَانَ	Abii Sufyaan	Abu Sufyan	10	10	75.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
6		Abi Sufyan			10	90.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
7	ابْنُ سَيِّهَابٍ	Ibnu Syihab	Ibnu Syihab	14	14	100.00%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
8		Ibnu Syihaabin			14	78.57%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
9	ابْنُ عَبَّاسٍ	Ibnu Abbas	Ibnu Abbas	19	19	100.00%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
10		Ibnu 'Abbaasin			0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						72.41%					80.00%	80.00%	80.00%	100.00%

Gambar 31. Pengujian pencocokkan 2 kata nama dengan batasan kemiripan 74%

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	الرُّهْرِيّ	Az Zuhri	Az Zuhri	17	17	100.00%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
2		Az-Zuhriy			17	88.89%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
3	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah	18	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
4		Abi Hurairah			18	91.67%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
5	أَبِي سُفْيَانَ	Abii Sufyaan	Abu Sufyan	10	10	75.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
6		Abi Sufyan			10	90.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
7	ابْنُ سَيِّهَابٍ	Ibnu Syihab	Ibnu Syihab	14	14	100.00%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
8		Ibnu Syihaabin			14	78.57%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
9	ابْنُ عَبَّاسٍ	Ibnu Abbas	Ibnu Abbas	19	19	100.00%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
10		Ibnu 'Abbaasin			0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						72.41%					80.00%	80.00%	80.00%	100.00%

Gambar 32. Pengujian pencocokkan 2 kata nama dengan batasan kemiripan 75%

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	الرُّهْرِيّ	Az Zuhri	Az Zuhri	17	17	100.00%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
2		Az-Zuhriy			17	88.89%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
3	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah	18	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
4		Abi Hurairah			18	91.67%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
5	أَبِي سُفْيَانَ	Abii Sufyaan	Abu Sufyan	10	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
6		Abi Sufyan			10	90.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
7	ابْنُ سَيِّهَابٍ	Ibnu Syihab	Ibnu Syihab	14	14	100.00%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
8		Ibnu Syihaabin			14	78.57%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
9	ابْنُ عَبَّاسٍ	Ibnu Abbas	Ibnu Abbas	19	19	100.00%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
10		Ibnu 'Abbaasin			0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						64.91%					70.00%	70.00%	70.00%	100.00%

Gambar 33. Pengujian pencocokkan 2 kata nama dengan batasan kemiripan 76%

No	Bahasa Arabnya	Query	Goldstandard	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	الرُّهْرِيّ	Az Zuhri	Az Zuhri	17	17	100.00%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
2		Az-Zuhriy			17	88.89%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
3	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah	18	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
4		Abi Hurairah			18	91.67%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
5	أَبِي سُفْيَانَ	Abii Sufyaan	Abu Sufyan	10	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
6		Abi Sufyan			10	90.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
7	ابْنُ سَيِّهَابٍ	Ibnu Syihab	Ibnu Syihab	14	14	100.00%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
8		Ibnu Syihaabin			14	78.57%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
9	ابْنُ عَبَّاسٍ	Ibnu Abbas	Ibnu Abbas	19	19	100.00%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
10		Ibnu 'Abbaasin			0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						64.91%					70.00%	70.00%	70.00%	100.00%

Gambar 34. Pengujian pencocokkan 2 kata nama dengan batasan kemiripan 77%

No	Bahasa Arabnya	Query	Goldstandar	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	الرُّهْرِيّ	Az Zuhri	Az Zuhri	17	17	100.00%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
2		Az-Zuhriy			17	88.89%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
3	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah	18	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
4		Abi Hurairah			18	91.67%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
5	أَبِي سُفْيَانَ	Abii Sufyaan	Abu Sufyan	10	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
6		Abi Sufyan			10	90.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
7	ابْنُ سَيِّهَابٍ	Ibnu Syihab	Ibnu Syihab	14	14	100.00%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
8		Ibnu Syihaabin			14	78.57%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
9	ابْنُ عَبَّاسٍ	Ibnu Abbas	Ibnu Abbas	19	19	100.00%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
10		Ibnu 'Abbaasin			0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						64.91%					70.00%	70.00%	70.00%	100.00%

Gambar 35. Pengujian pencocokkan 2 kata nama dengan batasan kemiripan 78%

No	Bahasa Arabnya	Query	Goldstandar	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	الرُّهْرِيّ	Az Zuhri	Az Zuhri	17	17	100.00%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
2		Az-Zuhriy			17	88.89%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
3	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah	18	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
4		Abi Hurairah			18	91.67%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
5	أَبِي سُفْيَانَ	Abii Sufyaan	Abu Sufyan	10	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
6		Abi Sufyan			10	90.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
7	ابْنُ سَيِّهَابٍ	Ibnu Syihab	Ibnu Syihab	14	14	100.00%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
8		Ibnu Syihaabin			0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
9	ابْنُ عَبَّاسٍ	Ibnu Abbas	Ibnu Abbas	19	19	100.00%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
10		Ibnu 'Abbaasin			0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						57.06%					60.00%	60.00%	60.00%	100.00%

Gambar 36. Pengujian pencocokkan 2 kata nama dengan batasan kemiripan 79%

No	Bahasa Arabnya	Query	Goldstandar	Jumlah di Gold Standard	Jumlah di Sistem	kemiripan	TP	FP	FN	TN	Precision	Recall	F-measure	Akurasi
1	الرُّهْرِيّ	Az Zuhri	Az Zuhri	17	17	100.00%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
2		Az-Zuhriy			17	88.89%	17	0	0	305	100.00%	100.00%	100.00%	100.00%
3	أَبِي هُرَيْرَةَ	Abii Hurairata	Abu Hurairah	18	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
4		Abi Hurairah			18	91.67%	18	0	0	304	100.00%	100.00%	100.00%	100.00%
5	أَبِي سُفْيَانَ	Abii Sufyaan	Abu Sufyan	10	0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
6		Abi Sufyan			10	90.00%	10	0	0	312	100.00%	100.00%	100.00%	100.00%
7	ابْنُ سَيِّهَابٍ	Ibnu Syihab	Ibnu Syihab	14	14	100.00%	14	0	0	308	100.00%	100.00%	100.00%	100.00%
8		Ibnu Syihaabin			0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
9	ابْنُ عَبَّاسٍ	Ibnu Abbas	Ibnu Abbas	19	19	100.00%	19	0	0	303	100.00%	100.00%	100.00%	100.00%
10		Ibnu 'Abbaasin			0	0.00%	0	0	0	322	0.00%	0.00%	0.00%	100.00%
Rata-rata nilai						57.06%					60.00%	60.00%	60.00%	100.00%

Gambar 37. Pengujian pencocokkan 2 kata nama dengan batasan kemiripan 80%

Nama :	Umur :	Asal :	1	2	3	4	5	6	7	8	9	10
Sania Bilqisty Aulia Rahman	18	Purwakarta Jawa Barat	Abuu Abbas	Abdullaah	A'isyah	Bilaa	Qotaadah	Muhammad	Umar	Syu'bah	Abdurrahman	
Nana	19	Banjarmasin	Abu Abbas	Abdullah	Aisyah	Bilal	Qotadah	Muhammad	Umar	Syu'bah	Abdurrohman	
Alifa	19	Jakarta selatan	Abu abbaas	abdulloh	aaisyah	Bilaa	Qotaadah	Muhammad	umar	Syu'bah	abdurrohman	
Rizki Nur Rohmah	19	Temanggung Jawa tengah	Abuu abbaasin	abdullohi	aaisyata	Bilaaun	Qotaadata	Mukhammadu	umaro	syu'batu	abdurrohmani	
Sensor	20	Klaten	Abu abbas	abdullah	aisyah	Bilal	Qotadah	Muhammad	umar	Syu'bah	abdureahman	
Maya	21	Yogyakarta	Abuu Abbaasin	Abekdullohi	Aaisyah	Bilaaun	Qotaadah	Mukhammadu	Umaro	Syu'bah	Abekdurrohmani	
Intan	19	Jogja	Abu Abbas	Abdulloh	aisyah	Bilal	Qotadah	Muhammad	umar	Syu'bah	abdurrohman	
Amalia Ilimi	19	Cirebon	Abu Abbaasin	Abdullahi	Aaisyata	Bilaaun	Qotaadatu	Muhammadu	Umaro	Syu'batu	Abdurrohmani	
Wahyu	20	Klaten	Abuu A'bbaas	A'bdullah	A'aaisyah	Bilaa	Qotaadah	Muhammad	U'mar	Syu'bah	A'bdurrohman	
Ina	21	Slleman	Abuu A'bbaas	A'b'dullah	A'a'aaisyah	Bilaa	Qataadah	Muhammad	U'mar	Syu'bah	A'b'durrohman	
Rosyad	19	Tangsel	Abu Abbas	Abdulloh	Aisyah	Bilal	Qotadah	Muhammad	Umar	Syu'bah	Abdurrohman	
Rosyad	19	Tangsel	Abu Abbas	Abdulloh	Aisyah	Bilal	Qotadah	Muhammad	Umar	Syu'bah	Abdurrohman	
Yahya	20	Yogyakarta	Abu Abbas	Abdullah	Aaisyah	Bilaa	Khatadah	Muhammad	Umar	Syu'bah	Abdurrohman	
Hilmi Lutfi	22	Majalengka	Abuu Abbaasin	Abdullah	Aaisyata	Bilaaun	Qataadata	Muhammadu	Umaro	Syu'batu	Abdurrahmani	
Alfian Yulianto	20	Sidoarjo	Abu Abbas	Abdullah	Aisyah	Bilal	Qotaadah	Muhammad	Umar	Syu'bah	Abdurrahman	
Rizal	23	Bandung	Abu Abbas	Abdullah	Aisyah	Bilal	Qatadah	Muhammad	Umar	Syu'bah	Abdurrahman	
Mep	21	Segawe	Abu Abbas	Abdulloh	Aisyah	Bilal	Qotadah	Muhammad	Umar	Syu'bah	Abdurrohman	
Muhammad Fauzan	24	Klaten	Abu Abbas	Abdullah	Aisyah	Bilal	Qatadah	Muhammad	Umar	Syu'bah	Abdurrahman	
Rinda	24	Bandung	Abu Abbas	Abdulloh	Aisyah	Bilal	Qotadah	Muhammad	Umar	Syubah	Abdurrohman	
Dian	23	Bandung	Abu Abbas	Abdullah	Aisyah	Bilal	Kotadat	Muhammad	Umar	Su'bat	Abdurrahman	
Aldi Abdul Aziz	24	Cimah	Abu A'bbaasin	Abdulloh	Aisyah	Bilaa	Qotaadah	Muhammad	Umar	Syu'bah	Abdurrahman	
Ziyan Novi M	21	Ponorogo	Abu Abbas	Abdullah	Aisyah	Bilal	Qotadah	Muhammad	Umar	Syu'bah	Abdurrohman	
AJI HARNANTO	24	BANDUNG	ABU ABBAS	ABDULLAH	AISYAH	BILAL	QOTADAH	MUHAMMAD	UMAR	SUKBAH	ABDURRAHMAN	

Gambar 38. Hasil kuesioner 1 kata nama dari responden bagian 1

KARTO	20	KEPANJEN	ABU	ABBAS	ABDULLAH	AISYAH	BILAL	QOTADAH	MUHAMMAD	UMAR	SYU'BAH	ABDURRAHMAN
Dandung Gusty Priyoga	18	Kebumen	Abu	Abbas	Abdullah	Aisyah	Bilal	Qatadah	Muhammad	Umar	Syu'bah	Abdurrahman
David	24	Purworejo	Abu	Abbasin	Abdulloh	Aisyah	Bilal	Khotadah	Muhammad	Umar	Syu'bah	Abdurrohman
Dafidl akmal	24	Bandung	Abuu	'Abbaasin	'Abdullahi	'Aaisyata	Bilaaalun	Qotaadata	Muchammadu	'Umara	Syu'batu	'Abdurrochmani
Dafidl akmal	24	Bandung	Abuu	'Abbaasin	'Abdullahi	'Aaisyata	Bilaaalun	Qotaadata	Muchammadu	'Umara	Syu'batu	'Abdurrochmani
Aisyah Maharani Harlan	21 thn	Pariaman	Abuu	A'bbaas	A'bdullah	Aa'isyah	Bilaaal	Qotaadah	Muhammad	U'mar	Syu'bah	A'bdurrahman
Moh Rivai	19	Indramayu	Abu	Abbaas	Abdulloh	Aaisyah	Bilaaal	Qotaadah	Mukhammad	Umar	Syu'bah	Abdurrohman
Enggar Tyas R	19	Jogja	Abuu	Abbaasin	Abdullah	Aaisyah	Bilaaal	Qotaadah	Muhammad	Umaro	Syu'bah	Abdurrohman
Okta	19	Jogja	Abuu	Abbas	Abdullah	Aaisyah	Bilaaal	Qataadah	Muhammad	Umar	Syu'bah	Abdurrahman
Aulia Riskafina Kusuma	19	Klaten	Abuu	abbaasin	abdullahi	aaisyata	bilaaalun	qataadata	muhammadu	umara	syu'batu	abdurrahmani
Dhanty amalia	19	Klaten	Abu	Abbaasi	Abdullah	Aaisyah	Bilal	Qatadah	Muhammad	Umar	Syu'bat	Abdurrahman
Lailatussa'adah	18	Bantul, Yogyakarta	Abuu	Abbaasi	Abdullah	Aisyah	Bikal	Khatadah	Muhammad	Umara	Syugbah	Abdurrahman
Aditya Bagus Wicaksono	20	Jepara, Jawa Tengah	Abu	Abbas	Abdulloh	Aisyah	Bilal	Khotadah	Mukhammad	Umar	Syu'bah	Abdurrohman
Aziz	19	Klaten	Abu	Abbas	Abdullah	Aisyah	Bilal	Qatadah	Muhammad	Umar	Syu'bat	Abdurrahman
Afif KH	23	Pati	Abu	Abas	Abdullah	Aisyah	Bilal	Kotadah	Muhammad	Umar	Syu'bah	Abdurrahman
Daffa Nisrina	18	Yogyakarta	Abu	Abbas	Abdullah	Aisyah	Bilal	Qotadah	Muhammad	Umar	Syu'bah	Abdurrahman
zilham	20	jawa	abuu	abbas	abdillah	aisyah	bilall	qotadah	muhammad	umar	syukbah	abdurrohman

Gambar 39. Hasil kuesioner 1 kata nama dari responden bagian 2

Nama	1	2	3	4	5
Dandung Gusty	Azzuhriyyi	Abii Hurairata	Abii Sufyaan	Ibnu Syihaabin	Ibnu 'Abbaasin
Syaeful	Az-Zuhriy	Abi Huroiroh	Abi Sufyan	Ibnu Syihab	Ibnu 'Abbas
Muhammad Fauzan	Az-zuhri	Abi Hurairah	Abi sufyan	Ibnu Syihab	Ibnu Abbas
Zalina Fatima Azzahra	Azzuhriy	Abi hurairah	Abi sufyan	Ibnu sihab	Ibnu 'abbas
Muhammad Farisqi Aziz	Az Zuhri	Abi Hurairah	Abi sufyaan	Ibnu Syihaab	Ibnu 'abaas
Nur Muthiah	Az-Zuhriyyi	Abi Hurairah	Abi Sufyan	Ibnu syihab	Ibnu Abbas

Gambar 40. Hasil kuesioner 2 kata nama dari responden