

BAB V

HASIL DAN PEMBAHASAN

5.1 Identifikasi Masalah

Penulis mengidentifikasi beberapa masalah yang akan dijadikan bahan penelitian :

1. Akurasi kemiripan data
2. Kecepatan dalam melakukan *screening* data
3. *Human error* saat melakukan *screening* data
4. Persentase kemiripan dari hasil perbandingan dua *string*

5.2 Dataset

Data yang digunakan dalam penelitian berjumlah 500 *record* data *dummy*, data *dummy* tersebut akan dibandingkan dengan data Daftar Terduga Teroris dan Organisasi Teroris (DTTOT), daftar Proliferasi (pengembang senjata pemusnah massal), daftar OFAC, PEP, data *dummy* dan daftar *blacklist* yang di dapatkan dari pihak Bank Indonesia dan PPATK (Pusat Pelaporan dan Analisis Transaksi Keuangan).

5.3 Environment

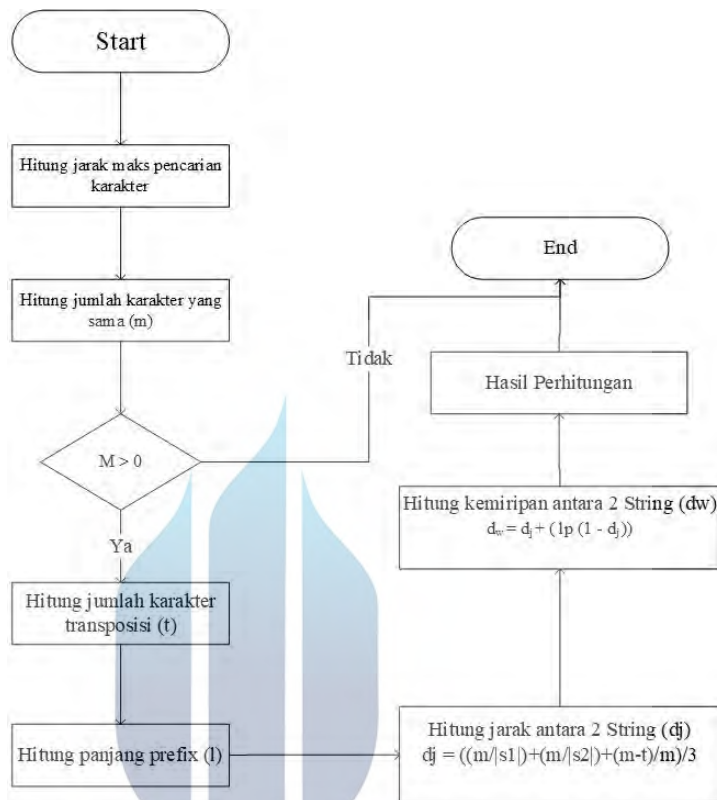
Penelitian ini dilakukan untuk mengetahui besar kemiripan antar data dengan data *blacklist reference*. Ada beberapa faktor yang mempengaruhi kemiripan data yaitu gelar, *case sensitive* sehingga besar kecilnya huruf dapat mempengaruhi jarak kedekatan, ada nya kata non alphabet seperti “'”. Untuk *threshold* kemiripan yang penulis tentukan adalah sebesar 90%.

5.4 Flowchart Algoritma

5.4.1 Algoritma Jaro-Winkler Distance

Algoritma Jaro Winkler distance menghitung kemiripan antara 2 string, algoritma jaro winkler distance memiliki 3 komponen dasar yaitu menghitung string atau panjang kata, menemukan jumlah karakter yang sama di dalam dua

string, mencari jumlah transposisi. Berikut flowchart algoritma jaro winkler distance.

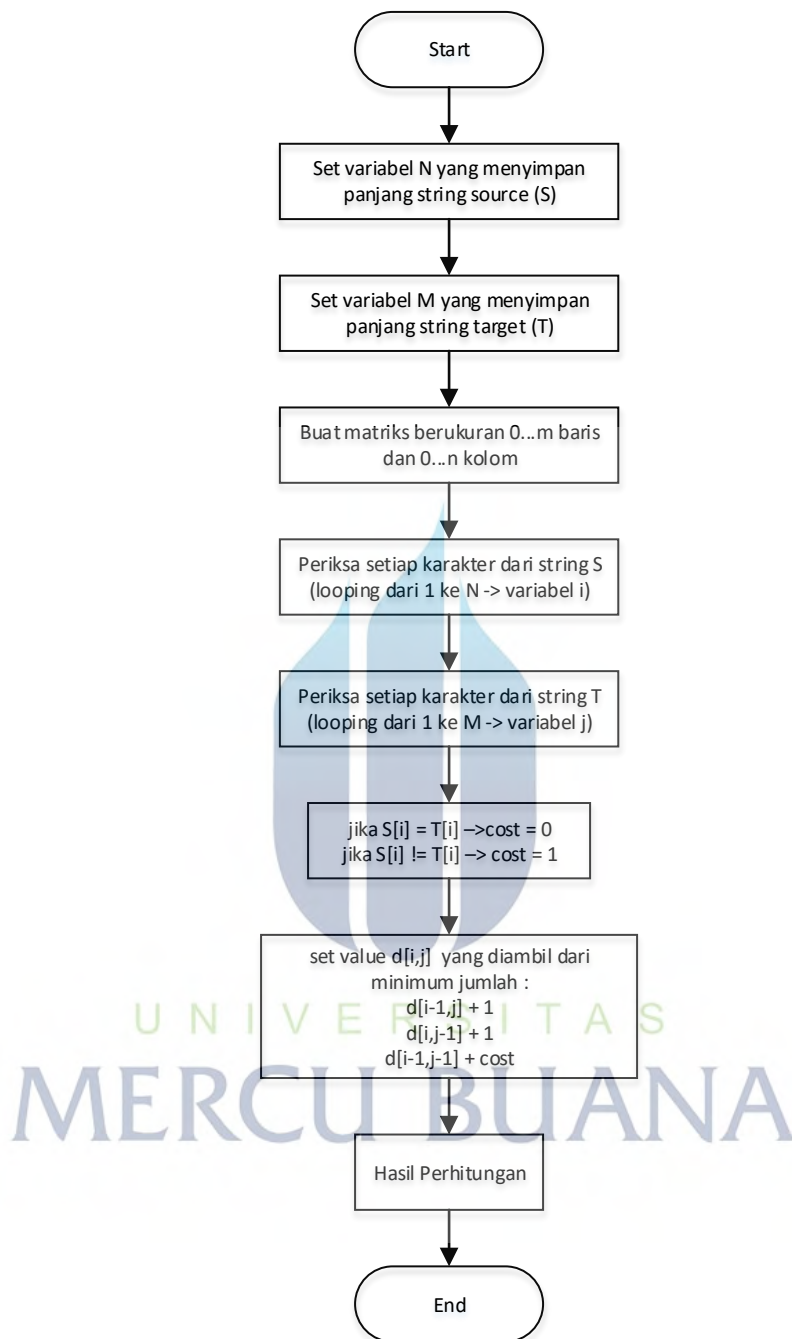


Gambar 5. 1 Flowchart algoritma jaro winkler distance

5.4.2 Algoritma Levenshtein Distance

Algoritma Levenshtein adalah sebuah matriks string yang digunakan untuk mengukur perbedaan atau jarak (distance) antara dua string. Nilai distance antara dua string ini ditentukan oleh jumlah minimum dari operasi-operasi perubahan yang diperlukan untuk melakukan transformasi dari suatu string menjadi string lainnya. Operasi-operasi tersebut adalah penyisipan (insertion), penghapusan (deletion), atau penukaran (substitution).

Berikut flowchart algoritma levenshtein distance :



Gambar 5. 2 Flowchart algoritma levensthein distance

5.5 Perhitungan Manual Algoritma

Berikut ini merupakan contoh perhitungan algoritma *Jaro-Winkler Distance* :

Contoh 1 :

1) S_1 : ABDULLAH = 8

S_2 : ABDULLOH = 8

		0	1	2	3	4	5	6	7
		A	B	D	U	L	L	A	H
0	A	√							
1	B		√						
2	D			√					
3	U				√				
4	L					√			
5	L						√		
6	O								
7	H								√

Dari table di atas dapat dilihat bahwa jumlah karakter yang sama persis yaitu $m=7$, pencocokan ABDULLaH dengan ABDULLoH memiliki urutan yang sama yaitu A-B-D-U-L-L-H, jadi tidak mengalami transposisi.

Maka nilai jaro *distance* adalah :

$$d_j = \frac{1}{3} \times \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right) = \frac{1}{3} \times \left(\frac{7}{8} + \frac{7}{8} + \frac{7-0}{7} \right) = 0,9166$$

Kemudian bila diperhatikan susunan s_1 dan s_2 dapat diketahui nilai $l = 4$, dan nilai konstan $p = 0,1$.

Maka nilai Jaro-Winkler *distance* :

$$d_w = d_j + (lp(1 - d_j)) = 0,9166 + (4 \times 0,1(1 - 0,9166)) = 0,94$$

Persentase Jaro *distance* manual adalah 91.6% dan persentase nilai Jaro-Winkler manual adalah 94 %. Sehingga kata ABDULLAH dan ABDULLOH dapat dikatakan mirip.

$$2) t = ABDULLOH = 8$$

$$s = ABDULLAH = 8$$

		A	B	D	U	L	L	O	H
	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
L	6	5	4	3	2	1	0	1	2
A	7	6	5	4	3	2	1	1	1
H	8	7	6	5	4	3	2	1	1

Dari table di atas dapat dilihat bahwa hasil algoritma levensthein = 1.

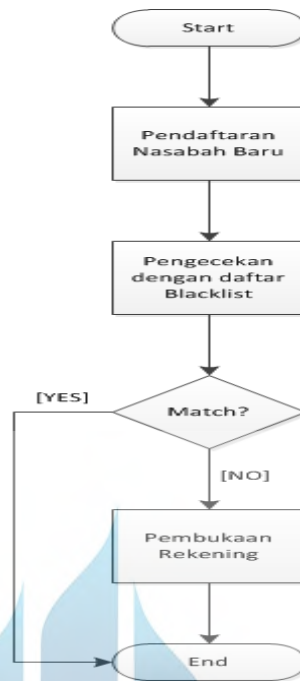
$$\text{Persentase Levensthein} = \frac{\text{jumlah kata yang sama}}{\text{jumlah kata target}} \times 100\%$$

$$= 7/8 \times 100 \%$$

$$= 87,5\%$$

5.6 Implementasi Algoritma Jaro Winkler *Distance* dan Levensthein *Distance*

5.6.1 Sistem Berjalan



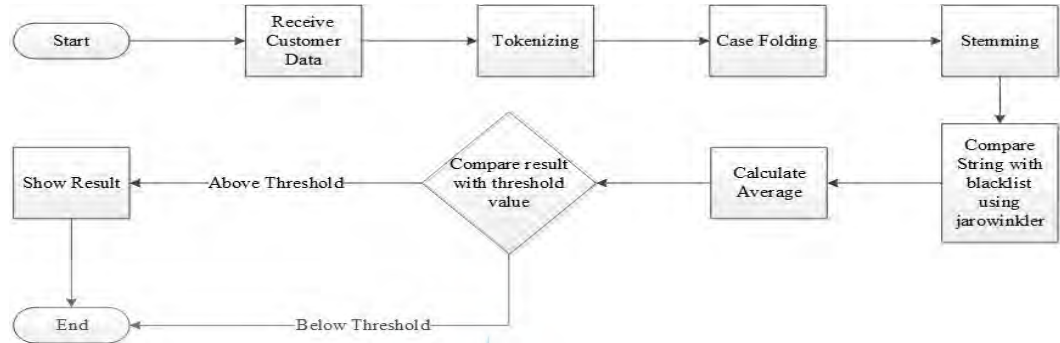
Gambar 5.3 Sistem Berjalan Screening Calon Nasabah

Gambar 5.3 *screening* calon nasabah dapat dilihat bahwa untuk melakukan pemeriksaan terhadap calon nasabah dilakukan secara manual dengan cara membandingkan data calon nasabah dengan daftar *blacklist* yang didapatkan dari PPATK yang berupa excel dan pdf, jika data calon nasabah cocok dengan data *blacklist* maka pihak bank menolak untuk melakukan pembukaan rekening, jika data calon nasabah tidak cocok dengan data *blacklist* maka pihak bank melakukan pembukaan rekening untuk calon nasabah.

5.6.2 Perancangan Sistem Usulan

Dalam perancangan sistem ini menjelaskan langkah-langkah pembuatan alur fungsi dan proses yang ada di sistem *screening* nasabah. Hal ini akan dibutuhkan untuk pedoman dalam membangun dan mengimplementasikan sistem.

Sebelum data string yang diinput oleh *user* dibandingkan dengan algoritma Jaro-Winkler *distance*, data *string* melalui proses *preprocessing* yaitu *tokenizing*, *case folding*, *stemming*.



Gambar 5.4 Flowchart Sistem

Proses *tokenizing/parsing* adalah tahap pemotongan *string input* berdasarkan tiap kata yang menyusunnya. Proses *tokenizing* yang penulis lakukan adalah dengan membagi nama menjadi kata, contoh nama “Regita Yulyonisfah” akan menjadi token “Regita” dan “Yulyonisfah” yang akan diproses pada tahapan berikutnya.

Proses *case folding* adalah mengubah token menjadi *lowercase*, karena algoritma jaro-winkler *case sensitive* sehingga besar kecilnya huruf dapat mempengaruhi jarak kedekatan. Karakter selain huruf dihilangkan dan dianggap *delimiter*. Contohnya token “Regita” akan diubah menjadi “regita” sedangkan token “yulyonisfah” tidak ada perubahan karena sudah seluruhnya huruf kecil.

Proses *stemming* dilakukan untuk mendapatkan *root word* dari sebuah token. Misalnya dengan menghilangkan karakter non alphabet dalam sebuah nama seperti “’”, menghapus token yang berisi gelar, mengubah string untuk meningkatkan kemiripan contohnya mengubah char “v” dengan “f”.

Selanjutnya setelah melalui *preprocessing* token tersebut dihitung jaraknya dengan algoritma jaro-winkler dan dilakukan proses penghitungan rata-rata jarak.

Dalam memodelkan proses yang terjadi pada rancangan sistem *screening* nasabah akan dibuat dalam bentuk model UML yaitu *activity diagram*, *use case diagram*, *class diagram*, *sequence diagram*.

5.6.2.1 Diagram Aktivitas (*Activity Diagram*)

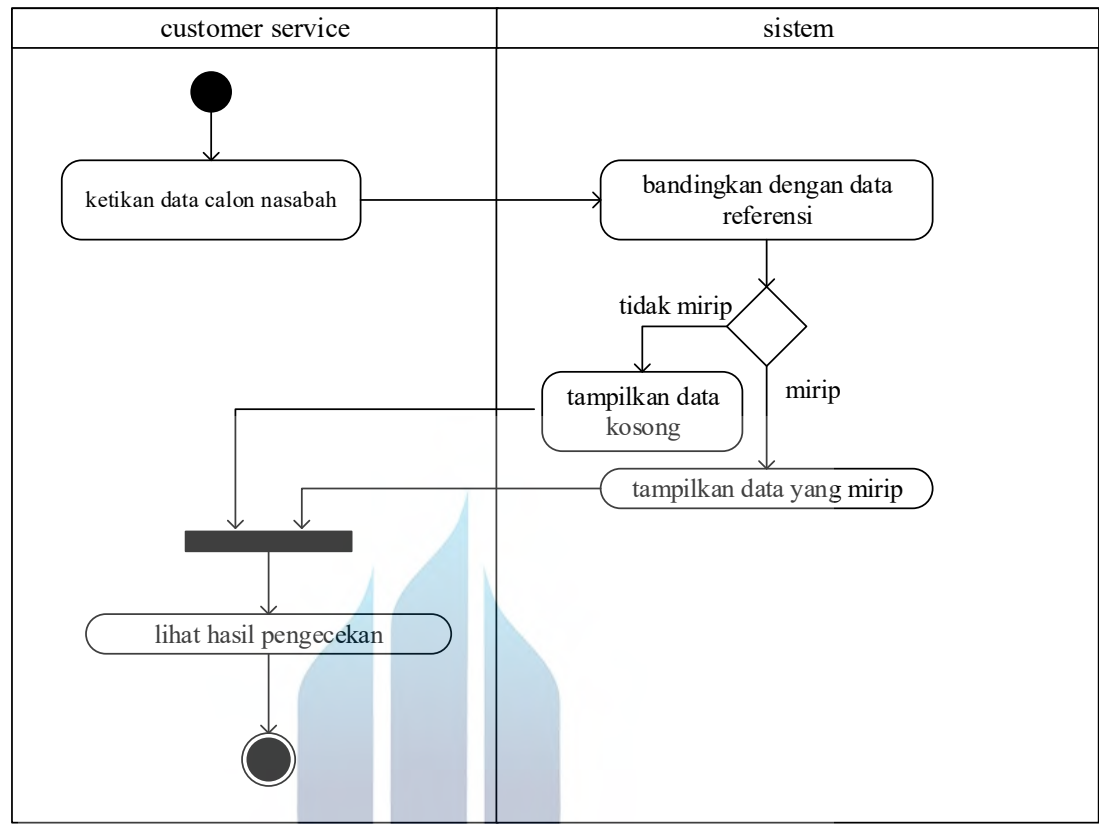
Activity diagram merupakan alur kerja pada setiap use case. Bagaimana masing masing alur berawal, *decision* yang mungkin terjadi, dan bagaimana sistem itu berakhir.

a. *Activity diagram screening new customer*

Aktivitas *input new customer* merupakan aktivitas pengecekan terhadap *customer* baru, yang mana data *customer* tersebut dibandingkan dengan data *blacklist reference* menggunakan algoritma jaro winkler *distance*. Jika data *customer* cocok dengan data *blacklist* maka dilakukan cetak surat pernyataan kemiripan, jika data nasabah tidak cocok dengan data *blacklist* maka dilakukan cetak surat pernyataan tidak ada kemiripan.



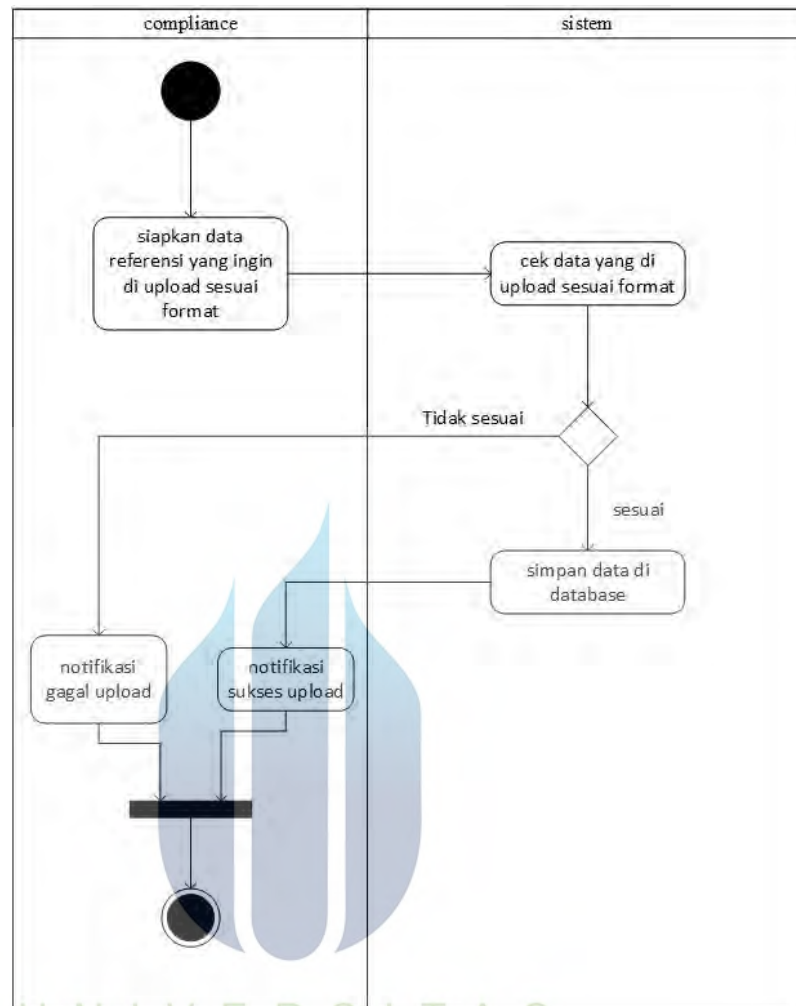
UNIVERSITAS
MERCU BUANA



Gambar 5.5 Activity diagram *screening new customer*

b. Activity diagram *upload blacklist reference*

Aktivitas yang dilakukan oleh *compliance*, *compliance* mengupload data *blacklist reference* seperti dhn, dtot, un, pep, data *blacklist reference* digunakan sebagai data pembanding.

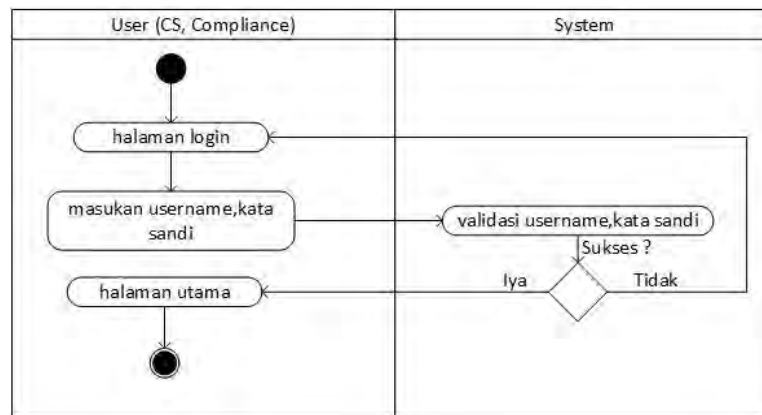


Gambar 5.6 Activity Upload Blacklist Reference

MERCU BUANA

c. Activity Diagram Login

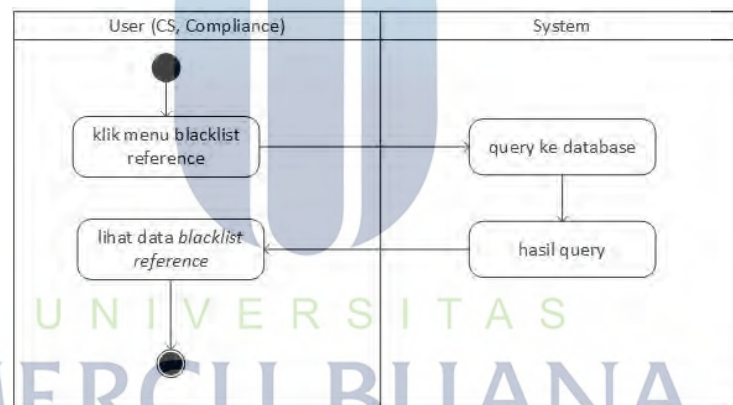
Aktivitas yang dilakukan oleh *user* sebelum menggunakan sistem *screening* nasabah. *User* harus *login* sebelum menggunakan sistem dengan menginputkan *username* dan *password* yang telah diberikan.



Gambar 5.7 Activity Diagram Login

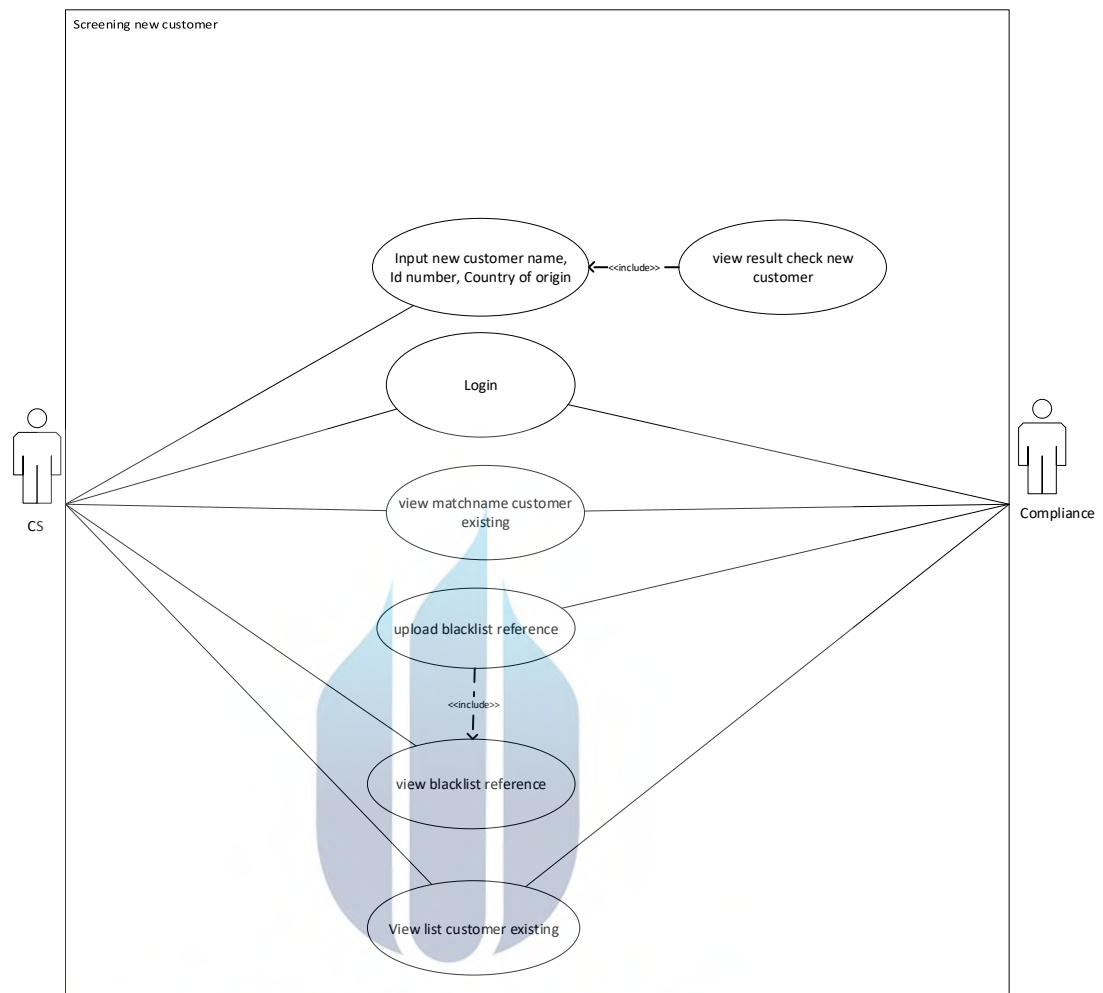
d. Activity Diagram *View Blacklist Reference*

Aktivitas yang dilakukan untuk melihat data *blacklist reference* yang telah di *upload*.

Gambar 5.8 Activity diagram *view blacklist reference*

5.6.2.2 Use case Diagram

Use case diagram merupakan suatu aktivitas yang menggambarkan urutan interaksi antar satu atau lebih aktor dan sistem. *Use case* yang akan dirancang yaitu *use case* diagram untuk pengaksesan melalui web. Adapun *use case* diagram untuk sistem *screening* nasabah dapat dilihat pada gambar berikut.



Gambar 5.9 Use case diagram screening nasabah
 Deskripsi pendefinisian aktor yang terlibat pada gambar-gambar *use case* diagram diatas dapat dilihat pada tabel 5.3 dibawah ini :

Tabel 5. 1 Deskripsi aktor use case

No	Aktor	Deskripsi
1	<i>Customer service</i>	Aktor yang bertugas melakukan pemeriksaan terhadap calon nasabah, mengajukan daftar blacklist.
2	<i>Compliance</i>	Aktor yang bertugas untuk memastikan perusahaan mematuhi peraturan pemerintah, mengupload data <i>blacklist</i> .

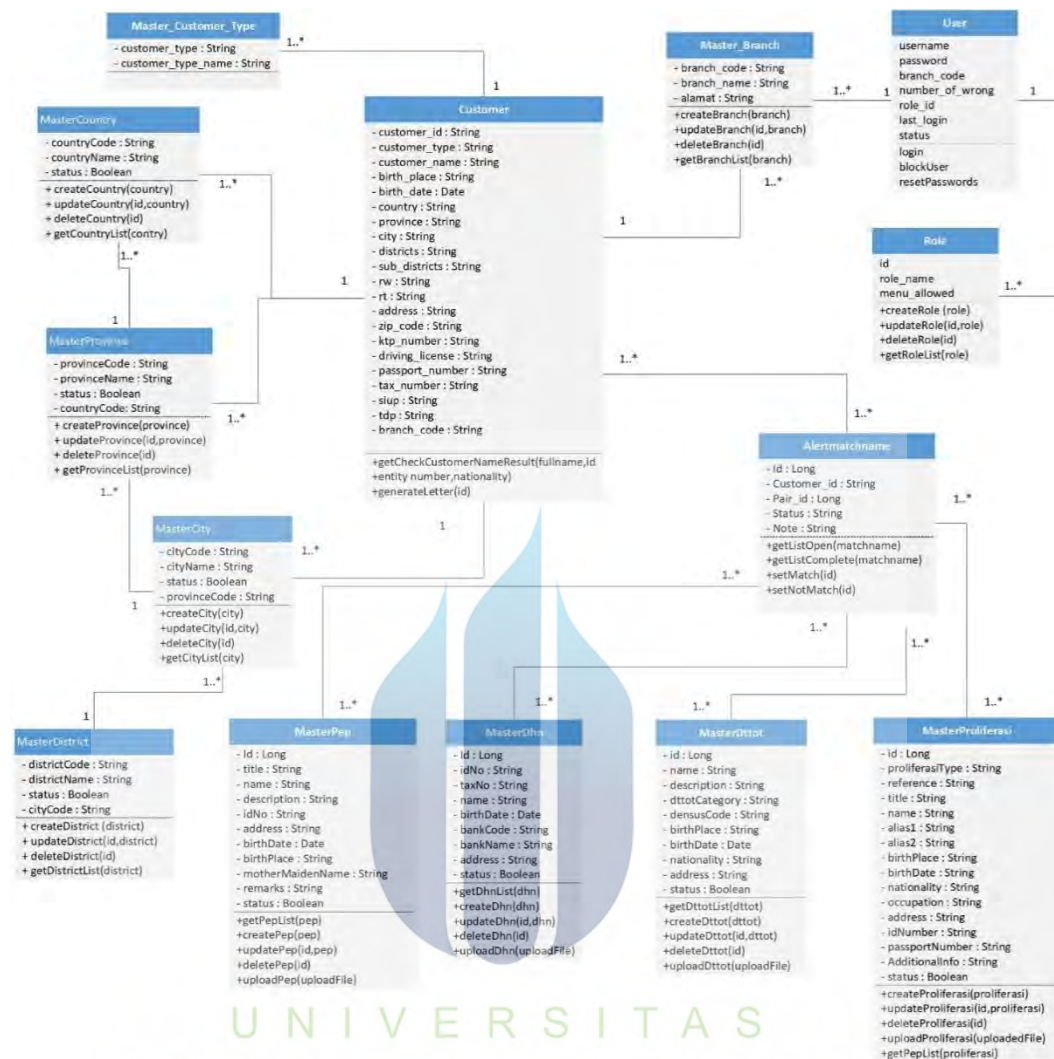
Identifikasi *use case* pada pemodelan *use case* diagram diatas untuk aplikasi ini dapat dilihat pada tabel 5.4 dibawah ini :

Tabel 5. 2 Deskripsi use case diagram

No	Use Case	Deskripsi
1	<i>Login</i>	<i>User</i> melakukan <i>login</i> ke aplikasi dengan memasukan <i>username</i> dan <i>password</i> .
2	<i>Input new customer name, id number, country of origin</i>	<i>User</i> (cs) menginputkan data (nama,nomor id, negara asal) <i>customer</i> pada sistem untuk dibandingkan dengan <i>blacklist reference</i> .
3	<i>View list customer existing</i>	<i>User</i> (cs, <i>compliance</i>) dapat melihat daftar data <i>customer existing</i> .
4	<i>View blacklist reference</i>	<i>User</i> (<i>compliance</i> dan CS) dapat melihat daftar <i>blacklist reference</i> .
4	<i>Upload blacklist reference</i>	<i>User</i> (<i>compliance</i>) dapat melakukan <i>upload</i> data <i>blacklist reference</i> .

5.6.2.3 Class Diagram

Class diagram menggambarkan struktur dan deskripsi *class*, *package*, dan objek yang saling terhubung. *Class diagram* dapat dilihat pada gambar 5.11.

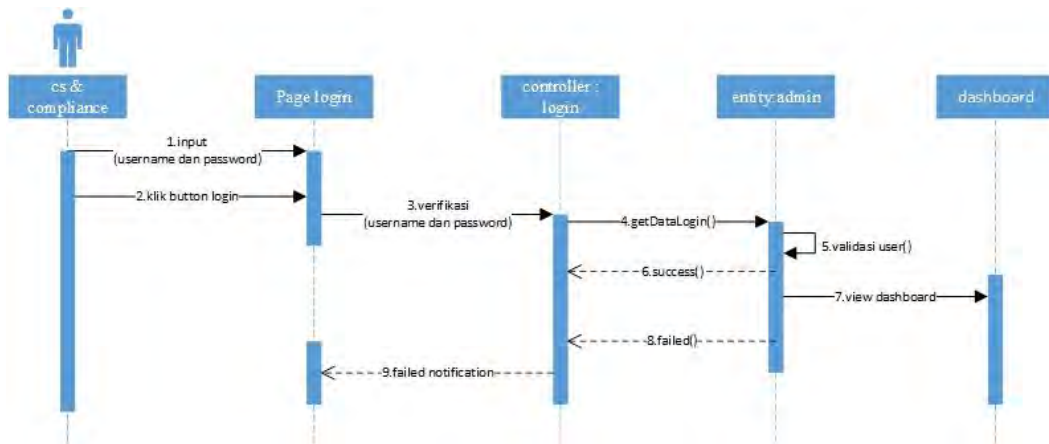


Gambar 5. 10 Class diagram

5.6.2.4 Sequence Diagram

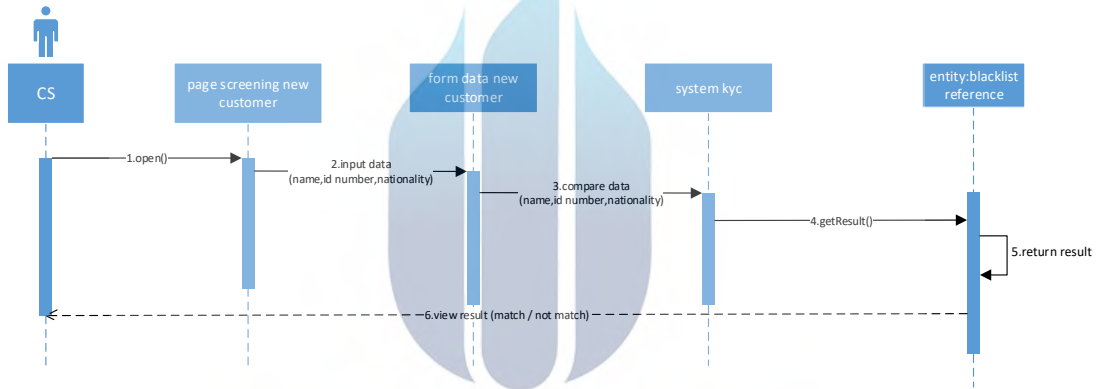
Sequence diagram menggambarkan interaksi antar objek di dalam dan di sekitar sistem (termasuk pengguna, *display* dan sebagainya) berupa pesan yang digambarkan terhadap waktu. Diagram ini juga menunjukkan serangkaian pesan yang dipertukarkan oleh objek-objek yang melakukan suatu tugas atau aksi tertentu.

a. *Sequence diagram login*



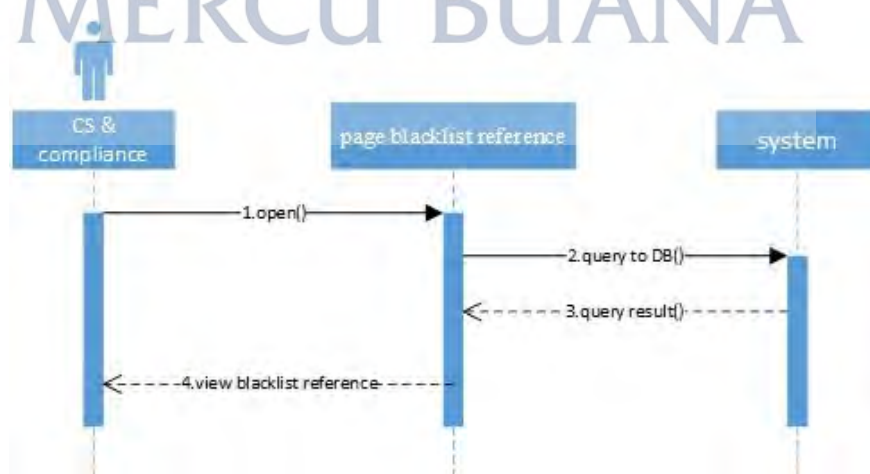
Gambar 5. 11 Sequence diagram login

b. *Sequence diagram screening new customer*



Gambar 5. 12 Sequence diagram screening new customer

c. *Sequence diagram view blacklist reference*



Gambar 5. 13 Sequence diagram view blacklist reference

5.6.3 Spesifikasi Database

Nama database : DBSCREENING

1. Nama tabel : customer

Primary key : customerid

Foreign key : branch_code, customer_type

Tabel 5. 3 tabel master customer

No	Fields	Type data	Length	Keterangan
1	customer_id	Varchar	16	Id customer
2	customer_type	Integer		Type customer
3	customer_name	Varchar	30	Nama customer
4	place_of_birth	Varchar	4	Tempat lahir
5	date_of_birth	Date		Tanggal lahir
6	Country	Varchar	3	Negara
7	Province	Varchar	2	Provinsi
8	City	Varchar	6	Kota/Kabupaten
9	Districts	Varchar	8	Kecamatan
10	sub_districts	Varchar	10	Kelurahan
11	Rw	Varchar	3	Rukun warga
12	Rt	Varchar	3	Rukun tetangga
13	Address	Varchar	100	Alamat
14	zip_code	Varchar	5	Kode pos
15	ktp_number	Varchar	16	Nomor KTP
16	driving_license	Varchar	12	Nomor SIM
17	passport_number	Varchar	10	Nomor passport
18	tax_number	Varchar	20	Nomor NPWP
19	Siup	Varchar	30	Surat izin usaha perdagangan
20	Tdp	Varchar	20	Tanda daftar perusahaan
21	branch_code	Varchar	6	Kode cabang

2. Nama tabel : master_branch

Primary key : branch_code

Foreign key : -

Tabel 5. 4 tabel master branch

No	Fields	Type data	Length	Keterangan
1	branch_code	Varchar	6	Kode cabang
2	branch_name	Varchar	15	Nama cabang
3	Alamat	Varchar	100	Alamat cabang

3. Nama tabel : master_customer_type

Primary key : customer_type

Foreign key : -

Tabel 5. 5 tabel master customer type

No	Fields	Type data	Length	Keterangan
1	customer_type	Integer	11	Tipe customer
2	customer_type_name	Varchar	10	Nama tipe customer

4. Nama tabel : tbl_alertmatchname

Primary key : id

Foreign key : -

Tabel 5. 6 tabel alert match name

No	Fields	Type data	Length	Keterangan
1	Id	Integer	11	Id matchname
2	customer_id	Varchar	16	Id customer
3	pair_id	Varchar	15	
4	Status	Integer	11	Status matchname

5	Note	Text		
---	------	------	--	--

5. Nama tabel : master_dhnlist

Primary key : serialno

Foreign key : -

Tabel 5. 7 tabel master dhnlist

No	Fields	Type data	Length	Keterangan
1	Id	Integer	11	Nomor serial
2	id_no	Varchar	32	Nomor blacklist
3	tax_no	Varchar	100	
4	Name	Varchar	100	Nama blacklist
5	Bankcode	Varchar	100	Kode bank
6	Bankname	Varchar	100	Nama bank
7	Address	Varchar	100	Alamat
8	Birthdate	Date		Tanggal lahir
9	Status	Integer	11	Status blacklist

6. Nama tabel : master_peplist

Primary key : id

Foreign key :

Tabel 5. 8 tabel master peplist

No	Fields	Type data	Length	Keterangan
1	Id	Integer	11	Id Pep
2	Name	Varchar	200	Nama pep
3	Title	Varchar	200	Gelar pep
4	Address	Varchar	250	Alamat 1
5	Dateofbirth	Date		Tanggal lahir
6	Placeofbirth	Varchar	100	Tempat lahir
7	Mothermaidenname	Varchar	100	

8	Remarks	Varchar	100	Keterangan
9	Status	Boolean		Pep Status Active/Inactive

7. Nama tabel : master_dttot

Primary key : id

Foreign key :

Tabel 5. 9 tabel master dttot

No	Fields	Type data	Length	Keterangan
1	Id	Integer	11	Id Dttot
2	Name	Varchar	300	Nama Dttot
3	Description	Varchar	300	Keterangan
4	dttot_category	Varchar	50	Kategori DTTOT
5	densus_code	Varchar	100	Kode dari densus
6	place_of_birth	Varchar	100	Tempat lahir
7	date_of_birth	Date		Tanggal lahir
8	Nationality	Varchar	30	Kewarganegaraan
9	Address	Varchar	100	Alamat
10	Status	Boolean		DTTOT Status Active/Inactive

8. Nama tabel : master_proliferasi

Primary key : id

Foreign key : -

Tabel 5. 10 tabel master proliferasi

No	Fields	Type data	Length	Keterangan
1	Id	Integer	11	Id Proliferasi
2	proliferasi_type	Varchar	25	Kategori proliferasi
3	Reference	Varchar	100	Kode proliferasi

4	Title	Varchar	30	Gelar proliferasi
5	Name	Varchar	500	Nama proliferasi
6	alias1	Varchar	500	Nama alias proliferasi
7	alias2	Varchar	500	Nama alias profliferasi
8	place_of_birth	Varchar	50	Tempat lahir
9	date_of_birth	Date		Tanggal lahir
10	Nationality	Varchar	50	Kewarganegaraan
11	Occupation	Varchar	50	Pekerjaan
12	Address	Varchar	100	Alamat
13	id_number	Varchar	50	No Id Proliferasi
14	passport_number	Varchar	50	No Passport Proliferasi
15	Additional_info	Varchar	100	Infomarsi tambahan
16	Status	Boolean		DTTOT Status Active/Inactive

9. Nama table : master_country

Primary key : country_code

Foreign_key :-

Tabel 5. 11 tabel master country

No	Fields	Type Data	Length	Keterangan
1	Country_code	Varchar	3	Kode Negara
2	Country_name	Varchar	50	Nama negara
3	Status	Boolean		Status Negara

10. Nama table : master_province

Primary key : province_code

Foreign_key : country_code

Tabel 5. 12 tabel master province

No	Fields	Type Data	Length	Keterangan
1	Country_code	Varchar	3	Kode Negara
2	Province_code	Bigint		Kode provinsi
3	province_name	Varchar	30	Nama provinsi
4	Status	Boolean		Status provinsi

11. Nama table : master_city

Primary key : city_code

Foreign_key : province_code

Tabel 5. 13 tabel master city

No	Fields	Type Data	Length	Keterangan
1	city_code	Varchar	4	Kode kota
2	Province_code	Varchar	2	Kode provinsi
3	city_name	Varchar	30	Nama kota
4	Status	Boolean		Status kota

12. Nama table : master_district

Primary key : district_code

Foreign_key : city_code

Tabel 5. 14 tabel master district

No	Fields	Type Data	Length	Keterangan
1	district_code	Varchar	10	Kode kecamatan
2	district_name	Varchar	6	nama kecamatan
3	City_code	Varchar	4	Kode kota
4	Status	Boolean		Status kecamatan

13. Nama tabel : User

Primary key : -

Foreign key : role_id , branch_code

Tabel 5. 15 tabel user

No	Fields	Type data	Length	Keterangan
1	Username	Varchar	15	Username
2	Password	Varchar	15	Password
3	Branch_code	Varchar	6	Kode cabang
4	Number_of_wrong	Integer	11	Jumlah kesalahan login
5	Role_id	Integer	11	Id role
6	Last_login	Datetime		Waktu terakhir login
7	Status	Integer	11	Status username

14. Nama tabel : Role

Primary key : id

Foreign key : -

Tabel 5. 16 tabel role

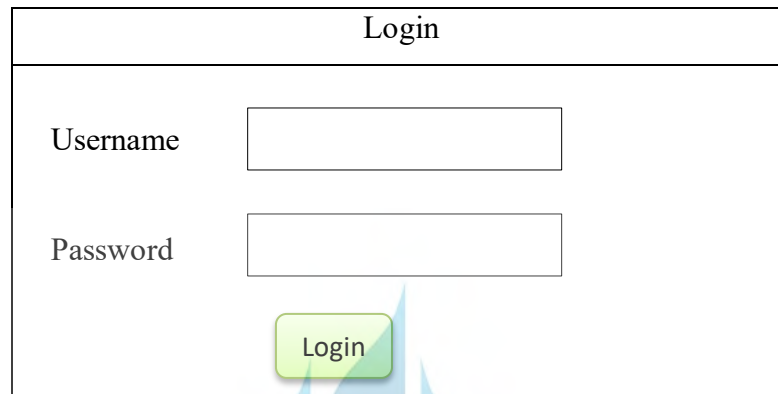
No	Fields	Type data	Length	Keterangan
1	Id	Integer	11	Id role
2	Role_name	Varchar	25	Nama role
3	Menu_allowed	Varchar	20	

5.6.4 Rancangan User Interface

Rancangan user interface dibuat secara umum, dimana rancangan user interface merupakan gambaran dari aplikasi yang akan di bangun.

5.6.4.1 Rancangan Halaman *Login*

Rancangan tampilan login adalah tampilan untuk menginputkan *username* dan *password* agar *user* dapat masuk kedalam aplikasi dan menggunakan aplikasi. Berikut merupakan rancangan tampilan login dapat dilihat pada gambar 5.16.

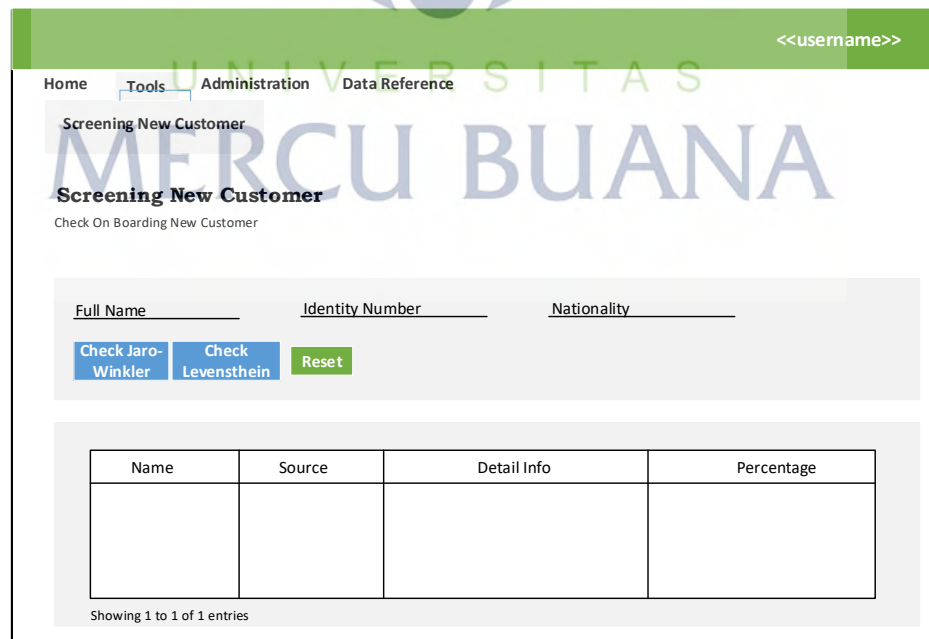


The image shows a simple login form titled "Login". It contains two input fields: "Username" and "Password". Below the "Password" field is a green "Login" button.

Gambar 5. 16 rancangan halaman login

5.6.4.2 Rancangan Halaman *Screening New Customer*

Halaman *screening new customer* merupakan halaman pengecekan calon nasabah baru yang mana data calon nasabah baru di bandingkan dengan *blacklist reference*.

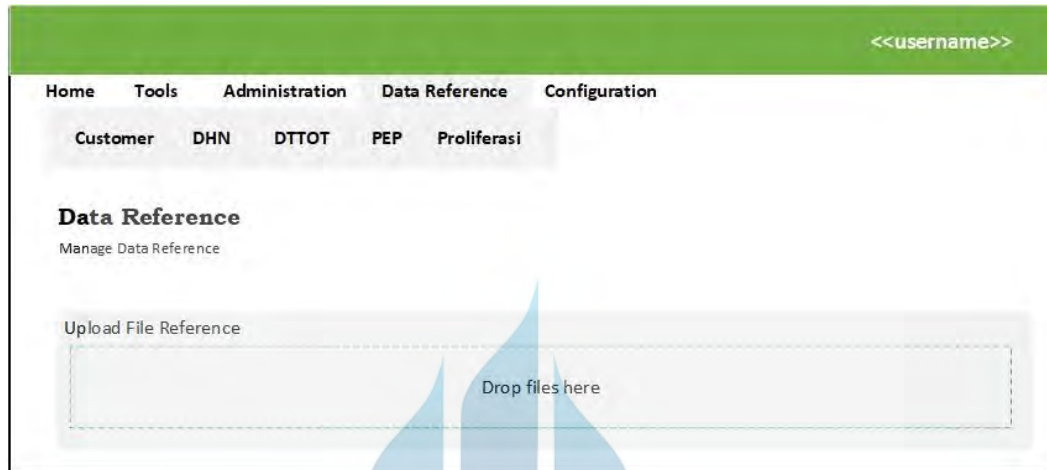


The image shows a screenshot of a web application interface for "Screening New Customer". The page has a green header with a user profile icon labeled "<<username>>". The main content area includes a navigation menu with "Home", "Tools", "Administration", and "Data Reference". The title "Screening New Customer" is displayed, along with the subtitle "Check On Boarding New Customer". Below this, there are three input fields: "Full Name", "Identity Number", and "Nationality". There are three buttons: "Check Jaro-Winkler" (blue), "Check Levensthein" (green), and "Reset" (green). At the bottom, there is a table with four columns: "Name", "Source", "Detail Info", and "Percentage". The table is currently empty. Below the table, it says "Showing 1 to 1 of 1 entries".

Gambar 5. 17 rancangan halaman screening new customer

5.6.4.3 Halaman *Input Upload Blacklist Reference*

Halaman ini merupakan tampilan untuk mengupload data *blacklist reference* (DHN, DTTOT, PEP, Proliferasi).



Gambar 5. 18 rancangan halaman input upload blacklist reference

5.6.4.4 Halaman *View Blacklist Reference*

Halaman *view blacklist reference* merupakan halaman yang berisikan list *blacklist reference* yang telah berhasil di *upload*.



Gambar 5. 19 halaman view blacklist reference

5.7 Evaluasi Kinerja Algoritma

Pada bagian ini, akan dilakukan beberapa pengujian dengan menginputkan nama pada menu *screening new customer* yang akan dibandingkan dengan data *blacklist reference* (DHN,DTTOT,PEP,Proliferasi) pada lampiran 1 yang sudah di upload pada sistem. Pengujian tersebut dilakukan dengan menghitung kemiripan data dengan menggunakan algoritma jaro winkler *distance* secara manual kemudian mencocokkan hasilnya dengan hasil dari sistem untuk memastikan bahwa sistem sudah menghasilkan aturan kemiripan data yang benar.

a) Pengujian dengan data inputan

Data yang akan dibandingkan dengan *blacklist reference* yaitu 'ZULKARNAIN'. Data yang diinputkan memiliki kemiripan dengan beberapa data yang ada di sistem yaitu data 'ZULKARNAEN' pada list DTTOT, data 'ZULKARNAINI' pada list DHN dan memiliki kemiripan 100 persen dengan data 'ZULKARNAIN' pada list DHN. Berikut perhitungan manual algoritma jaro-winkler *distance* data tersebut.

- $S_1 = \text{ZULKARNAIN} = 10$
 $S_2 = \text{ZULKARNAEN} = 10$

		0	1	2	3	4	5	6	7	8	9
		Z	U	L	K	A	R	N	A	I	N
0	Z	√									
1	U		√								
2	L			√							
3	K				√						
4	A					√					
5	R						√				
6	N							√			
7	A								√		
8	E										
9	N										√

Dari tabel diatas dapat dilihat bahwa jumlah karakter yang sama persis yaitu $m = 9$, pencocokan ZULKARNAIN dan ZULKARNAEN memiliki urutan yang sama yaitu Z-U-L-K-A-R-N-A-N, jadi tidak memiliki transposisi.

Maka nilai jaro *distance* adalah :

$$d_j = \frac{1}{3} \times \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right) = \frac{1}{3} \times \left(\frac{9}{10} + \frac{9}{10} + \frac{9-0}{9} \right) = 0,9333$$

Kemudian bila diperhatikan susunan s_1 dan s_2 dapat diketahui nilai $l = 4$, dan nilai konstan $p = 0,1$.

Maka nilai Jaro-Winkler *distance* :

$$d_w = d_j + (lp(1 - d_j)) = 0,9333 + (4 \times 0,1(1 - 0,9333)) = 0,96$$

Persentase Jaro *distance* manual adalah 93,3 % dan persentase nilai Jaro-Winkler manual adalah 96 %.

Hasil perbandingan nama yang diinputkan ZULKARNAIN memiliki kemiripan dengan data ZULKARNAEN adalah 96,01%. Hasil perbandingan dari perhitungan manual dan sistem adalah sama. Dari perbandingan tersebut, dapat dikatakan bahwa sistem sudah menjalankan perhitungan kemiripan untuk data nama yang diinputkan.

- $S_1 = \text{ZULKARNAIN} = 10$

$$S_2 = \text{ZULKARNAINI} = 11$$

		0	1	2	3	4	5	6	7	8	9
	Z	U	L	K	A	R	N	A	I	N	
0	Z	√									
1	U		√								
2	L			√							
3	K				√						
4	A					√					
5	R						√				
6	N							√			
7	A								√		

8	I									√	
9	N										√
10	I									√	

Dari tabel diatas dapat dilihat bahwa jumlah karakter yang sama persis yaitu $m = 10$, pencocokan ZULKARNAIN dan ZULKARNAINi memiliki urutan yang sama yaitu Z-U-L-K-A-R-N-A-I-N, jadi tidak memiliki transposisi.

Maka nilai jaro *distance* :

$$d_j = \frac{1}{3} \times \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right) = \frac{1}{3} \times \left(\frac{10}{10} + \frac{10}{11} + \frac{10-0}{10} \right) = 0.9696$$

Kemudian bila diperhatikan susunan s_1 dan s_2 dapat diketahui nilai $l = 4$, dan nilai konstan $p = 0,1$.

$$d_w = d_j + (lp(1 - d_j)) = 0,9696 + (4 \times 0,1(1 - 0,9696)) = 0,9817$$

Persentase Jaro *distance* manual adalah 96,8% dan persentase nilai Jaro-Winkler manual adalah 98,17%.

Hasil perbandingan nama yang diinputkan ZULKARNAIN memiliki kemiripan dengan data ZULKARNAINI adalah 98,19%. Hasil perbandingan dari perhitungan manual dan sistem adalah sama. Dari perbandingan tersebut, dapat dikatakan bahwa sistem sudah menjalankan perhitungan kemiripan untuk data nama yang diinputkan.

Gambar 5. 20 hasil pengujian dengan menginputkan nama menggunakan algoritma jaro-winkler

- $S_1 = \text{THORIQAN} = 8$
 $S_2 = \text{THORIQUN} = 8$

		0	1	2	3	4	5	6	7
		T	H	O	R	I	Q	A	N
0	T	√							
1	H		√						
2	O			√					
3	R				√				
4	I					√			
5	Q						√		
6	U								
7	N								√

Dari tabel diatas dapat dilihat bahwa jumlah karakter yang sama persis yaitu $m = 9$, pencocokan THORIQAN dan THORIQUN memiliki urutan yang sama yaitu T-H-O-R-I-Q-N, jadi tidak memiliki transposisi.

Maka nilai jaro *distance* adalah :

$$d_j = \frac{1}{3} \times \left(\frac{m}{|s_1|} + \frac{m}{|s_2|} + \frac{m-t}{m} \right) = \frac{1}{3} \times \left(\frac{7}{8} + \frac{7}{8} + \frac{7-0}{7} \right) = 0,91666$$

Kemudian bila diperhatikan susunan s_1 dan s_2 dapat diketahui nilai $l = 4$, dan nilai konstan $p = 0,1$.

Maka nilai Jaro-Winkler *distance* :

$$d_w = d_j + (lp(1 - d_j)) = 0,9166 + (4 \times 0,1(1 - 0,9166)) = 0,949$$

Persentase Jaro *distance* manual adalah 91 % dan persentase nilai Jaro-Winkler manual adalah 94 %.

- $S_1 = \text{TATANK} = 6$
 $S_2 = \text{TATANG} = 6$

		0	1	2	3	4	5
		T	A	T	A	N	K
0	T	√					
1	A		√				
2	T			√			
3	A				√		
4	N					√	
5	G						

Dari tabel diatas dapat dilihat bahwa jumlah karakter yang sama persis yaitu $m = 5$, pencocokan TATANK dan TATANG memiliki urutan yang sama yaitu T-A-T-A-N , jadi tidak memiliki transposisi.

Maka nilai jaro *distance* adalah :

$$d_j = \frac{1}{3} \times \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right) = \frac{1}{3} \times \left(\frac{5}{6} + \frac{5}{6} + \frac{5-0}{5} \right) = 0,8888$$

Kemudian bila diperhatikan susunan s_1 dan s_2 dapat diketahui nilai $l = 4$, dan nilai konstan $p = 0,1$.

Maka nilai Jaro-Winkler *distance* :

$$d_w = d_j + (lp(1 - d_j)) = 0,88888 + (4 \times 0,1(1 - 0,8888)) = 0,9336$$

Persentase Jaro *distance* manual adalah 88 % dan persentase nilai Jaro-Winkler manual adalah 93 %.

- $S_1 = \text{FAHIR} = 5$
 $S_2 = \text{FIHIR} = 5$

		0	1	2	3	4
		F	A	H	I	R
0	F	√				
1	I					
2	H			√		
3	I				√	
4	R					√

Dari tabel diatas dapat dilihat bahwa jumlah karakter yang sama persis yaitu $m = 5$, pencocokan FAHIR dan FIHIR memiliki urutan yang sama yaitu F-H-I-R, jadi tidak memiliki transposisi.

Maka nilai jaro *distance* adalah :

$$d_j = \frac{1}{3} \times \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right) = \frac{1}{3} \times \left(\frac{4}{5} + \frac{4}{5} + \frac{4-0}{4} \right) = 0,8666$$

Kemudian bila diperhatikan susunan s_1 dan s_2 dapat diketahui nilai $l = 4$, dan nilai konstan $p = 0,1$.

Maka nilai Jaro-Winkler *distance* :

$$d_w = d_j + (lp(1 - d_j)) = 0,8666 + (1 \times 0,1(1 - 0,8666)) = 0,8816$$

Persentase Jaro *distance* manual adalah 86 % dan persentase nilai Jaro-Winkler manual adalah 88 %.

- $S_1 = \text{AZIZAROHMAN} = 11$
 $S_2 = \text{AZIZIRHAMAN} = 11$

		0	1	2	3	4	5	6	7	8	9	10
		A	Z	I	Z	A	R	O	H	M	A	N
0	A	√										
1	Z		√									
2	I			√								
3	Z				√							
4	I											
5	R						√					
6	A											
7	H								√			
8	M									√		
9	A										√	
10	N											√

Dari tabel diatas dapat dilihat bahwa jumlah karakter yang sama persis yaitu $m = 9$, pencocokan THORIQAN dan THORIQUN memiliki urutan yang sama yaitu A-Z-I-Z-R-A-H-M-A-N , jadi tidak memiliki transposisi.

Maka nilai jaro *distance* adalah :

$$d_j = \frac{1}{3} \times \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right) = \frac{1}{3} \times \left(\frac{9}{11} + \frac{9}{11} + \frac{9-0}{9} \right) = 0,8787$$

Kemudian bila diperhatikan susunan s_1 dan s_2 dapat diketahui nilai $l = 4$, dan nilai konstan $p = 0,1$.

Maka nilai Jaro-Winkler *distance* :

$$d_w = d_j + (lp(1 - d_j)) = 0,8787 + (4 \times 0,1(1 - 0,8787)) = 0,9272$$

Persentase Jaro *distance* manual adalah 87 % dan persentase nilai Jaro-Winkler manual adalah 92 %.

- $S_1 = \text{SHYAMS} = 6$
 $S_2 = \text{SHAMS} = 5$

		0	1	2	3	4	5
		S	H	Y	A	M	S
0	S	√					
1	H		√				
2	A						
3	M						
4	S						

Dari tabel diatas dapat dilihat bahwa jumlah karakter yang sama persis yaitu $m = 5$, pencocokan SHYAMS dan SHAMS memiliki urutan yang sama yaitu S-H-A-M-S, jadi tidak memiliki transposisi.

Maka nilai jaro *distance* adalah :

$$d_j = \frac{1}{3} \times \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right) = \frac{1}{3} \times \left(\frac{5}{6} + \frac{5}{5} + \frac{5-0}{5} \right) = 0,9444$$

Kemudian bila diperhatikan susunan s_1 dan s_2 dapat diketahui nilai $l = 2$, dan nilai konstan $p = 0,1$.

Maka nilai Jaro-Winkler *distance* :

$$d_w = d_j + (lp(1 - d_j)) = 0,9444 + (2 \times 0,1(1 - 0,8787)) = 0,95$$

Persentase Jaro *distance* manual adalah 94 % dan persentase nilai Jaro-Winkler manual adalah 95 %.

- $S_1 = \text{HAMDULAH} = 8$
 $S_2 = \text{HAMDULLAH} = 9$

		0	1	2	3	4	5	6	7
		H	A	M	D	U	L	A	H
0	H	√							
1	A		√						
2	M			√					
3	D				√				
4	U								
5	L						√		
6									
7									√
8									
9									
10									

Dari tabel diatas dapat dilihat bahwa jumlah karakter yang sama persis yaitu $m = 8$, pencocokan HAMDULAH dan HAMDULLAH memiliki urutan yang sama yaitu H-A-M-D-U-L-A-H, jadi tidak memiliki transposisi.

Maka nilai jaro *distance* adalah :

$$d_j = \frac{1}{3} \times \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right) = \frac{1}{3} \times \left(\frac{8}{8} + \frac{8}{9} + \frac{8-0}{8} \right) = 0,9629$$

Kemudian bila diperhatikan susunan s_1 dan s_2 dapat diketahui nilai $l = 4$, dan nilai konstan $p = 0,1$.

Maka nilai Jaro-Winkler *distance* :

$$d_w = d_j + (lp(1 - d_j)) = 0,9629 + (4 \times 0,1(1 - 0,9629)) = 0,9777$$

Persentase Jaro *distance* manual adalah 96 % dan persentase nilai Jaro-Winkler manual adalah 97 %.

- $S_1 = \text{FHAIZ} = 5$
 $S_2 = \text{FAIZ} = 4$

		0	1	2	3	4
		F	H	A	I	Z
0	F	√				
1	A					
2	I					
3	Z					

Dari tabel diatas dapat dilihat bahwa jumlah karakter yang sama persis yaitu $m = 4$, pencocokan FHAIZ dan FAIZ memiliki urutan yang sama F-A-I-Z, jadi tidak memiliki transposisi.

Maka nilai jaro *distance* adalah :

$$d_j = \frac{1}{3} \times \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right) = \frac{1}{3} \times \left(\frac{4}{5} + \frac{4}{4} + \frac{4-0}{4} \right) = 0,93333$$

Kemudian bila diperhatikan susunan s_1 dan s_2 dapat diketahui nilai $l = 1$, dan nilai konstan $p = 0,1$.

Maka nilai Jaro-Winkler *distance* :

$$d_w = d_j + (lp(1 - d_j)) = 0,9333 + (1 \times 0,1(1 - 0,9333)) = 0,94$$

Persentase Jaro *distance* manual adalah 93 % dan persentase nilai Jaro-Winkler manual adalah 94 %.

- $S_1 = \text{FAREDSAL} = 8$
 $S_2 = \text{FAREDSAAL} = 9$

		0	1	2	3	4			
		F	A	R	E	D	S	A	L
0	F	√							
1	A		√						
2	R			√					
3	E				√				
4	D					√			
5	S						√		
6	A								
7	A								
8	L								√

Dari tabel diatas dapat dilihat bahwa jumlah karakter yang sama persis yaitu $m = 8$, pencocokan FAREDSAL dan FAREDSAAL memiliki urutan yang sama F-A-R-E-D-S-A-L, jadi tidak memiliki transposisi.

Maka nilai jaro *distance* adalah :

$$d_j = \frac{1}{3} \times \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right) = \frac{1}{3} \times \left(\frac{8}{8} + \frac{8}{9} + \frac{8-0}{8} \right) = 0,9629$$

Kemudian bila diperhatikan susunan s_1 dan s_2 dapat diketahui nilai $l = 4$, dan nilai konstan $p = 0,1$.

Maka nilai Jaro-Winkler *distance* :

$$d_w = d_j + (lp(1 - d_j)) = 0,9629 + (4 \times 0,1(1 - 0,0148)) = 0,9777$$

Persentase Jaro *distance* manual adalah 96% dan persentase nilai Jaro-Winkler manual adalah 97 %.

Berikut perhitungan manual algoritma levensthein *distance* data tersebut.

- $t = \text{ZULKARNAIN} = 10$
 $s = \text{ZULKARNAEN} = 10$

		Z	U	L	K	A	R	N	A	I	N
	0	1	2	3	4	5	6	7	8	9	10
Z	1	0	1	2	3	4	5	6	7	8	9
U	2	1	0	1	2	3	4	5	6	7	8
L	3	2	3	0	1	2	3	4	5	6	7
K	4	3	4	3	0	1	2	3	4	5	6
A	5	4	5	4	3	0	1	2	3	4	5
R	6	5	6	5	4	3	0	1	2	3	4
N	7	6	7	6	5	4	3	0	1	2	3
A	8	7	8	7	6	5	4	3	0	1	2
E	9	8	9	8	7	6	5	4	3	1	1
N	10	9	8	9	8	7	6	0	4	3	1

Dari table di atas dapat dilihat bahwa hasil algoritma levensthein = 1.

$$\text{Persentase Levensthein} = \frac{\text{jumlah kata yang sama}}{\text{jumlah kata target}} \times 100\%$$

$$= 9/10 \times 100 \%$$

$$= 90\%$$



Gambar 5. 21 hasil pengujian dengan menginputkan nama menggunakan algoritma levensthein

- t = THORIQAN = 8
- s = THORIQAN = 8

	A	B	D	U	L	L	O	H	
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
L	6	5	4	3	2	1	0	1	2
A	7	6	5	4	3	2	1	1	1
H	8	7	6	5	4	3	2	1	1

Dari table di atas dapat dilihat bahwa hasil algoritma Levenshtein = 1.

$$\text{Persentase Levenshtein} = \frac{\text{jumlah kata yang sama}}{\text{jumlah kata target}} \times 100\%$$

$$= 7/8 \times 100 \%$$

$$= 87,5\%$$

- t = ABDULLOH = 8
s = ABDULLAH = 8

		A	B	D	U	L	L	O	H
	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
L	6	5	4	3	2	1	0	1	2
A	7	6	5	4	3	2	1	1	1
H	8	7	6	5	4	3	2	1	1

Dari table di atas dapat dilihat bahwa hasil algoritma Levenshtein = 1.

$$\text{Persentase Levenshtein} = \frac{\text{jumlah kata yang sama}}{\text{jumlah kata target}} \times 100\%$$

$$= 7/8 \times 100 \%$$

$$= 87,5\%$$

- t = TATANG = 6
s = TATANK = 6

		T	A	T	A	N	G
	0	1	2	3	4	5	6
T	1	0	1	2	3	4	5
A	2	1	0	1	2	3	4
T	3	2	1	0	1	2	3
A	4	3	2	1	0	1	2
N	5	4	3	2	1	0	1
K	6	5	4	3	2	1	1

Dari table di atas dapat dilihat bahwa hasil algoritma levensthein = 1.

$$\text{Persentase Levensthein} = \frac{\text{jumlah kata yang sama}}{\text{jumlah kata target}} \times 100\%$$

$$= 5/6 \times 100 \%$$

$$= 83,3\%$$

- t = FIHIR = 5
s = FAHIR = 5

		F	I	H	I	R
	0	1	2	3	4	5
F	1	0	1	2	3	4
A	2	1	0	1	2	3
H	3	2	1	0	1	2
I	4	3	2	1	0	1
R	5	4	3	2	1	1

Dari table di atas dapat dilihat bahwa hasil algoritma levensthein = 1.

$$\text{Persentase Levensthein} = \frac{\text{jumlah kata yang sama}}{\text{jumlah kata target}} \times 100\%$$

$$= 4/5 \times 100 \%$$

$$= 80\%$$

- $t = \text{AZIZIRAHMAN} = 11$
 $s = \text{AZIZAROHMAN} = 11$

		A	Z	I	Z	I	R	A	H	M	A	N
	0	1	2	3	4	5	6	7	8	9	10	11
A	1	0	1	2	3	4	5	6	7	8	9	10
Z	2	1	0	1	2	3	4	5	6	7	8	9
I	3	2	1	0	1	2	3	4	5	6	7	8
Z	4	3	2	1	0	1	2	3	4	5	6	7
A	5	4	3	2	1	1	1	2	3	4	5	6
R	6	5	4	3	2	1	0	1	2	3	4	5
O	7	6	5	4	3	2	1	1	1	2	3	4
H	8	7	6	5	4	3	2	2	0	1	2	3
M	9	8	7	6	5	4	3	3	1	0	1	2
A	10	9	8	7	6	5	4	4	2	1	0	1
N	11	10	9	8	7	6	5	5	3	2	1	2

Dari table di atas dapat dilihat bahwa hasil algoritma levensthein = 1.

$$\text{Persentase Levensthein} = \frac{\text{jumlah kata yang sama}}{\text{jumlah kata t arg et}} \times 100\%$$

$$= 9/11 \times 100 \%$$

$$= 81,8\%$$

- $t = \text{SHAMS} = 5$
 $s = \text{SHYAMS} = 6$

		S	H	A	M	S
	0	1	2	3	4	5
S	1	0	1	2	3	4
H	2	1	0	1	2	3

Y	3	2	1	1	1	2
A	4	3	2	1	1	1
M	5	4	3	2	1	1
S	6	5	4	3	2	1

Dari table di atas dapat dilihat bahwa hasil algoritma levensthein = 1.

$$\text{Persentase Levensthein} = \frac{\text{jumlah kata yang sama}}{\text{jumlah kata target}} \times 100\%$$

$$= 5/6 \times 100 \%$$

$$= 83,3\%$$

- t = HAMDULLAH = 9
- s = HAMDULAH = 8

		H	A	M	D	U	L	L	A	H
	0	1	2	3	4	5	6	7	8	9
H	1	0	1	2	3	4	5	6	7	8
A	2	1	0	1	2	3	4	5	6	7
M	3	2	1	0	1	2	3	4	5	6
D	4	3	2	1	0	1	2	3	4	5
U	5	4	3	2	1	0	1	2	3	4
L	6	5	4	3	2	1	0	1	2	3
A	7	6	5	4	3	2	1	1	1	2
H	8	5	4	3	4	3	2	1	1	1

Dari table di atas dapat dilihat bahwa hasil algoritma levensthein = 1.

$$\text{Persentase Levensthein} = \frac{\text{jumlah kata yang sama}}{\text{jumlah kata target}} \times 100\%$$

$$= 8/9 \times 100 \%$$

$$= 88,8\%$$

- $t = \text{FAIZ} = 4$
 $s = \text{FHAIZ} = 5$

		F	A	I	Z
	0	1	2	3	4
F	1	0	1	2	3
H	2	1	1	1	2
A	3	2	1	1	1
I	4	3	2	1	1
Z	5	4	3	2	1

Dari table di atas dapat dilihat bahwa hasil algoritma levensthein = 1.

$$\text{Persentase Levensthein} = \frac{\text{jumlah kata yang sama}}{\text{jumlah kata target}} \times 100\%$$

$$= 4/5 \times 100 \%$$

$$= 80\%$$

- $t = \text{FAREDSAAL} = 9$
 $s = \text{FAREDSAL} = 8$

		F	A	R	E	D	S	A	A	L
	0	1	2	3	4	5	6	7	8	9
F	1	0	1	2	3	4	5	6	7	8
A	2	1	0	1	2	3	4	5	6	7
R	3	2	1	0	1	2	3	4	5	6
E	4	3	2	1	0	1	2	3	4	5
D	5	4	3	2	1	0	1	2	3	4
S	6	5	4	3	2	1	0	1	2	3
A	7	6	5	4	3	2	1	0	1	2
L	8	7	5	5	4	3	2	1	1	1

Dari table di atas dapat dilihat bahwa hasil algoritma levensthein = 1.

$$\text{Persentase Levensthein} = \frac{\text{jumlah kata yang sama}}{\text{jumlah kata target}} \times 100\%$$

$$= 8/9 \times 100 \%$$

$$= 88,8\%$$

