

BAB IV. ANALISIS DAN PERANCANGAN SISTEM

4.1 Deskripsi Sistem

Pada skripsi ini dirancang dan dibangun sebuah aplikasi pendeteksi plagiarisme dokumen teks lintas bahasa menggunakan metode *Winnowing*. Aplikasi ini dibuat berbasis web sehingga dapat diakses oleh *user* dengan mudah.

Pengguna aplikasi ini terdiri dari dua macam pengguna yaitu *User* dan *Admin*. *Admin* adalah pengguna yang memiliki hak akses untuk melihat tahapan proses perhitungan mulai dari *preprocessing* sampai dengan hasil persentase *Jaccard Coefficient*. Sedangkan *user* hanya dapat melihat tahapan perhitungan dari hasil persentase *Jaccard Coefficient*.

Selain itu, *user* juga dapat memasukkan dokumen berupa inputan teks maupun *file* PDF, menghapus dokumen teks, melakukan perhitungan dan menghapus daftar riwayat perhitungan. Pada saat memasukkan dokumen berupa *file* PDF, *user* dapat memasukkan lima *file* PDF sekaligus dan memberi judul masing-masing tiap *file* PDF.

Tahapan proses yang dijalankan dalam pendeteksian plagiarisme dokumen teks lintas bahasa yaitu sebagai berikut:

A. Input Dokumen

Pada Tabel 4.1 berikut memperlihatkan isi dokumen yang akan diuji untuk mencari tingkat plagiarismenya menggunakan metode *Winnowing*. Dokumen dapat berisi teks berbahasa Indonesia maupun bahasa Inggris.

Tabel 4. 1 Isi Dokumen Uji dan Dokumen Pembanding

Dokumen Uji	Dokumen Pembanding
The cross-language plagiarism is an illegal use of content from one language to another language.	Plagiarisme lintas bahasa adalah menyalin suatu karya milik orang lain dari bahasa yang berbeda dan diterjemahkan ke bahasa lain.

B. Menerjemahkan Dokumen Teks

Pada Tabel 4.2 merupakan hasil terjemahan dari bahasa Inggris ke bahasa Indonesia. Dokumen yang berbahasa Inggris akan langsung diterjemahkan ke dalam bahasa Indonesia oleh sistem.

Tabel 4. 2 Proses Terjemahan Dokumen

Dokumen Uji	Dokumen Pembanding
Plagiarisme lintas bahasa adalah penggunaan konten secara illegal dari satu bahasa ke bahasa lain.	Plagiarisme lintas bahasa adalah menyalin suatu karya milik orang lain dari bahasa yang berbeda dan diterjemahkan ke bahasa lain.

C. Pembatasan Uji Dokumen

Pada proses ini dilakukan pembatasan isi dokumen dengan menghilangkan teks sebelum abstrak dan setelah daftar pustaka. Pembatasan kutipan dan sitasi juga dilakukan pada tahap ini.

D. *Preprocessing*

Pada Tabel 4.3 merupakan proses *preprocessing* di mana terdapat 3 tahap yaitu *case folding*, *tokenizing* dan *filtering*.

Tabel 4. 3 Tahap *Preprocessing*

Dokumen Uji	Dokumen Pembanding
plagiarismelintsbahasapenggun aankontenillegalbahasabahasa	plagiarismelintsbahasamenyalinkaryamilik orangbahasaberbedaditerjemahkanbahasa

E. Membentuk Nilai *N-Gram*

Pada Tabel 4.4 merupakan hasil pembentukan *n-gram* dengan nilai *n* adalah 8. Dalam pemilihan nilai *n* didasarkan pada pengujian manual pada Tabel 5.5 pada Sub Bab Analisis Rencana Pengujian di mana semakin besar nilai *n-gram* dan *w-gram* yang digunakan maka hasil persentase kemiripannya akan semakin kecil.

Tabel 4. 4 Contoh Hasil *N-Gram* dengan $n = 9$

Dokumen Uji	Dokumen Pembanding
plagiarislagiarism giarisme giarismel iarismeli arismelin rismelint ismelinta smelintas melintasb elintasba lintasbah	plagiarislagiarism giarisme giarismel iarismeli arismelin rismelint ismelinta smelintas melintasb elintasba lintasbah

Dokumen Uji			Dokumen Pembanding			
intasbaha	ntasbahas	tasbahasa	intasbaha	ntasbahas	tasbahasa	
asbahasap	sbahasape	bahasapen	asbahasam	sbahasame	bahasamen	
ahasapeng	hasapengg	asapenggu	ahasameny	hasamenya	asamenyal	
sapenggun	apengguna	penggunaa	samenyali	amenyalin	menyalink	
enggunaan	nggunaank	ggunaanko	enyalinka	nyalinkar	yalinkary	
gunaankon	unaankont	naankonte	alinkarya	linkaryam	inkaryami	
aankonten	ankonteni	nkontenil	nkaryamil	karyamili	aryamilik	
kontenile	ontenileg	ntenilega	tenilegal	ryamiliko	yamilikor	amilikora
enilegalb	nilegalba	ilegalbah	legalbaha	milikoran	ilikorang	likorangb
egalbahas	galbahasa	albahasab	ikorangba	ikorangba	korangbah	orangbaha
lbahasaba	bahasabah	ahasabaha	rangbahas	angbahasa	ngbahasab	
hasabahas	asabahasa		gbahasabe	bahasaber	ahasaberb	
			hasaberbe	asaberbed	saberbeda	
			aberbedad	berbedadi	erbedadit	
			rbedadite	bedaditer	edaditerj	daditerje
			aditerjem	diterjema	iterjemah	
			terjemahk	erjemahka	rjemahkan	
			jemahanb	emahanba	mahkanbah	
			ahkanbaha	hkanbahas	kanbahasa	

F. Membentuk Nilai *Hash/Rolling Hash*

Pada Tabel 4.5 ditampilkan hasil perhitungan nilai *hash*. Di bawah ini merupakan contoh perhitungan mencari nilai *rolling hash* menggunakan rumus pada Persamaan 2.1 di atas.

$$\begin{aligned}
 H(\text{plagiaris}) &= (112 \times 11^{(9-1)}) + (108 \times 11^{(9-2)}) + (97 \times 11^{(9-3)}) \\
 &\quad + (103 \times 11^{(9-4)}) + (105 \times 11^{(9-5)}) + (97 \times 11^{(9-6)}) \\
 &\quad + (114 \times 11^{(9-7)}) + (105 \times 11^{9-8}) + (115 \times 11^{9-9}) \\
 &= 24008194672 + 2104614468 + 171841417 + 16588253 \\
 &\quad + 1537305 + 129107 + 13794 + 1155 + 115 \\
 &= 26302920286
 \end{aligned}$$

Pada perhitungan di atas, nilai 112, 108, 97, 103, 105, 97, 114, 105 dan 115 merupakan kode ASCII dari karakter huruf *p*, *l*, *a*, *g*, *i*, *a*, *r*, *i* dan *s*. Kode ASCII dapat dilihat pada Tabel 3.1 diatas. Nilai 9 merupakan banyak karakter atau panjang rangkaian *n-gram*. Sedangkan nilai 11 merupakan basis yang digunakan. Hasil nilai *hash* dari kata *plagiari* adalah 26302920286.

Tabel 4. 5 Hasil *Rolling Hash*

Dokumen Uji		Dokumen Pembanding	
26302920286	25241981863	26302920286	25241981863
23003449966	24317023707	23003449966	24317023707
24618648709	23220628354	24618648709	23220628354
26705985983	24959809136	26705985983	24959809136
26973393056	25543339249	26973393056	25543339249
23960433517	25412052000	23960433517	25412052000
24874221469	26031928719	24874221469	26031928719
26976969996	23224737912	26976969996	23224737909
26751191106	23099117811	26751191073	23099117448
23011422306	24404719442	23011418331	24404675711
23225354115	26757969348	23224873065	26752677793
23173678460	26189537130	23115471368	25549259128
23994767148	25789721944	24025552186	26128357369
24312695485	24571038272	28037685170	23102866356
27412808935	25661018539	25410603998	24858293455
22896958029	23145612397	25856720558	25049680233
25880810448	25314669019	23246079733	26985951147
26160956375	26038326521	28039425957	23122015013
27047345829	23998872061	25621239226	24817333270
25834875977	24809389841	25406158513	24809393112
25318780793	23848238210	25318816781	26206581751
24177903616	23088327701	26540205675	23136225748
25250678781	23099116067	25777557299	24178884380
23011403116	24404508364	23099116121	23011403711
23223032237		24404514895	23223104081

Dokumen Uji	Dokumen Pemanding	
	26733218961	22901424206
	23194740344	24063270182
	26543255312	23169771772
	23788615880	23522057990
	22947868899	23705631959
	24967182553	27054500635
	24077574926	26700607505
	24900645879	23964649520
	25458428033	23026410141
	24569585639	25038882262

G. Membentuk *Window*

Pada Tabel 4.6 merupakan hasil pembentukan *window* dengan nilai *window* adalah 10. *Window* pertama berisi nilai *hash* pertama sampai dengan nilai *hash* ke 10. *Window* kedua berasal dari nilai *hash* kedua sampai nilai *hash* ke 11 dan dilanjutkan sampai *window* terakhir. Penentuan nilai *w-gram* adalah 10 didasarkan dari hasil perhitungan terbaik antara perhitungan *Winnowing* dengan hasil dari aplikasi Plagiarsm Checker X seperti pada Tabel 5.5 pada Sub Bab Analisis Rencana Pengujian. *Window* pertama dari dokumen uji adalah 26302920286, 25241981863, 23003449966, 24317023707, 24618648709, 23220628354 dan 26705985983. Sedangkan *window* kedua diambil dari nilai *hash* kedua yaitu 25241981863, jadi *window* kedua yaitu 25241981863, 23003449966, 24317023707, 24618648709, 23220628354, 26705985983 dan 24959809136.

Tabel 4. 6 Hasil Pembentukan *Window*

Dokumen Uji		Dokumen Pemanding	
{26302920286	25241981863	{26302920286	25241981863
23003449966	24317023707	23003449966	24317023707
24618648709	23220628354	24618648709	23220628354
26705985983},	{25241981863	26705985983},	{25241981863
23003449966	24317023707	23003449966	24317023707
24618648709	23220628354	24618648709	23220628354

Dokumen Uji		Dokumen Pemandang	
26705985983	24959809136},	26705985983	24959809136},
{23003449966	24317023707	{23003449966	24317023707
24618648709	23220628354	24618648709	23220628354
26705985983	24959809136	26705985983	24959809136
26973393056},	{24317023707	26973393056},	{24317023707
24618648709	23220628354	24618648709	23220628354
26705985983	24959809136	26705985983	24959809136
26973393056	25543339249},	26973393056	25543339249},
{24618648709	23220628354	{24618648709	23220628354
26705985983	24959809136	26705985983	24959809136
26973393056	25543339249	26973393056	25543339249
23960433517},	{23220628354	23960433517},	{23220628354
26705985983	24959809136	26705985983	24959809136
26973393056	25543339249	26973393056	25543339249
23960433517	25412052000},	23960433517	25412052000},
{26705985983	24959809136	{26705985983	24959809136
26973393056	25543339249	26973393056	25543339249
23960433517	25412052000	23960433517	25412052000
24874221469},	{24959809136	24874221469},	{24959809136
26973393056	25543339249	26973393056	25543339249
23960433517	25412052000	23960433517	25412052000
24874221469	26031928719},	24874221469	26031928710},
{26973393056	25543339249	{26973393056	25543339249
23960433517	25412052000	23960433517	25412052000
24874221469	26031928719	24874221469	26031928719
26976969996},	{25543339249	26976969996},	{25543339249
23960433517	25412052000	23960433517	25412052000
24874221469	26031928719	24874221469	26031928719
26976969996	23224737912},	26976969996	23224737909},
{23960433517	25412052000	{23960433517	25412052000
24874221469	26031928719	24874221469	26031928719

Dokumen Uji		Dokumen Pemandangan	
26976969996	23224737912	26976969996	23224737909
26751191106},	{25412052000	26751191073},	{25412052000
24874221469	26031928719	24874221469	26031928719
26976969996	23224737912	26976969996	23224737909
26751191106	23099117811},	26751191073	23099117448},
{24874221469	26031928719	{24874221469	26031928719
26976969996	23224737912	26976969996	23224737909
26751191106	23099117811	26751191073	23099117448
23011422306},	{26031928719	23011418331},	{26031928719
26976969996	23224737912	26976969996	23224737909
26751191106	23099117811	26751191073	23099117448
23011422306	24404719442},	23011418331	24404675711},
{26976969996	23224737912	{26976969996	23224737909
26751191106	23099117811	26751191073	23099117448
23011422306	24404719442	23011418331	24404675711
23225354115},	{23224737912	23224873065},	{23224737909
26751191106	23099117811	26751191073	23099117448
23011422306	24404719442	23011418331	24404675711
23225354115	26757969348},	23224873065	26752677793},
{26751191106	23099117811	{26751191073	23099117448
23011422306	24404719442	23011418331	24404675711
23225354115	26757969348	23224873065	26752677793
23173678460},	{23099117811	23115471368},	{23099117448
23011422306	24404719442	23011418331	24404675711
23225354115	26757969348	23224873065	26752677793
23173678460	26189537130},	23115471368	25549259128},
{23011422306	24404719442	{23011418331	24404675711
23225354115	26757969348	23224873065	26752677793
23173678460	26189537130	23115471368	25549259128
23994767148},	{24404719442	24025552186},	{24404675711
23225354115	26757969348	23224873065	26752677793

Dokumen Uji		Dokumen Pemandangan	
23173678460	26189537130	23115471368	25549259128
23994767148	25789721944},	24025552186	26128357369},
{23225354115	26757969348	{23224873065	26752677793
23173678460	26189537130	23115471368	25549259128
23994767148	25789721944	24025552186	26128357369
24312695485},	{26757969348	28037685170},	{26752677793
23173678460	26189537130	23115471368	25549259128
23994767148	25789721944	24025552186	26128357369
24312695485	24571038272},	28037685170	23102866356},
{23173678460	26189537130	{23115471368	25549259128
23994767148	25789721944	24025552186	26128357369
24312695485	24571038272	28037685170	23102866356
27412808935},	{26189537130	25410603998},	{25549259128
23994767148	25789721944	24025552186	26128357369
24312695485	24571038272	28037685170	23102866356
27412808935	25661018539},	25410603998	24858293455},
{23994767148	25789721944	{24025552186	26128357369
24312695485	24571038272	28037685170	23102866356
27412808935	25661018539	25410603998	24858293455
22896958029},	{25789721944	25856720558},	{26128357369
24312695485	24571038272	28037685170	23102866356
27412808935	25661018539	25410603998	24858293455
22896958029	23145612397},	25856720558	25049680233},
{24312695485	24571038272	{28037685170	23102866356
27412808935	25661018539	25410603998	24858293455
22896958029	23145612397	25856720558	25049680233
25880810448},	{24571038272	23246079733},	{23102866356
27412808935	25661018539	25410603998	24858293455
22896958029	23145612397	25856720558	25049680233
25880810448	25314669019},	23246079733	26985951147},
{27412808935	25661018539	{25410603998	24858293455

Dokumen Uji		Dokumen Pemandangan	
22896958029	23145612397	25856720558	25049680233
25880810448	25314669019	23246079733	26985951147
26160956375},	{25661018539	28039425957},	{24858293455
22896958029	23145612397	25856720558	25049680233
25880810448	25314669019	23246079733	26985951147
26160956375	26038326521},	28039425957	23122015013},
{22896958029	23145612397	{25856720558	25049680233
25880810448	25314669019	23246079733	26985951147
26160956375	26038326521	28039425957	23122015013
27047345829},	{23145612397	25621239226},	{25049680233
25880810448	25314669019	23246079733	26985951147
26160956375	26038326521	28039425957	23122015013
27047345829	23998872061},	25621239226	24817333270},
{25880810448	25314669019	{23246079733	26985951147
26160956375	26038326521	28039425957	23122015013
27047345829	23998872061	25621239226	24817333270
25834875977},	{25314669019	25406158513},	{26985951147
26160956375	26038326521	28039425957	23122015013
27047345829	23998872061	25621239226	24817333270
25834875977	24809389841},	25406158513	24809393112},
{26160956375	26038326521	{28039425957	23122015013
27047345829	23998872061	25621239226	24817333270
25834875977	24809389841	25406158513	24809393112
25318780793},	{26038326521	25318816781},	{23122015013
27047345829	23998872061	25621239226	24817333270
25834875977	24809389841	25406158513	24809393112
25318780793	23848238210},	25318816781	26206581751},
{27047345829	23998872061	{25621239226	24817333270
25834875977	24809389841	25406158513	24809393112
25318780793	23848238210	25318816781	26206581751
24177903616},	{23998872061	26540205675},	{24817333270

Dokumen Uji		Dokumen Pemandangan	
25834875977	24809389841	25406158513	24809393112
25318780793	23848238210	25318816781	26206581751
24177903616	23088327701},	26540205675	23136225748},
{25834875977	24809389841	{25406158513	24809393112
25318780793	23848238210	25318816781	26206581751
24177903616	23088327701	26540205675	23136225748
25250678781},	{24809389841	25777557299},	{24809393112
25318780793	23848238210	25318816781	26206581751
24177903616	23088327701	26540205675	23136225748
25250678781	23099116067},	25777557299	24178884380},
{25318780793	23848238210	{25318816781	26206581751
24177903616	23088327701	26540205675	23136225748
25250678781	23099116067	25777557299	24178884380
23011403116},	{23848238210	23099116121},	{26206581751
24177903616	23088327701	26540205675	23136225748
25250678781	23099116067	25777557299	24178884380
23011403116	24404508364},	23099116121	23011403711},
{24177903616	23088327701	{26540205675	23136225748
25250678781	23099116067	25777557299	24178884380
23011403116	24404508364	23099116121	23011403711
23223032237}		24404514895},	{23136225748
		25777557299	24178884380
		23099116121	23011403711
		24404514895	23223104081},
		{25777557299	24178884380
		23099116121	23011403711
		24404514895	23223104081
		26733218961},	{24178884380
		23099116121	23011403711
		24404514895	23223104081
		26733218961	22901424206},

Dokumen Uji	Dokumen Pemandang
	{23099116121 23011403711
	24404514895 23223104081
	26733218961 22901424206
	23194740344}, {23011403711
	24404514895 23223104081
	26733218961 22901424206
	23194740344 24063270182},
	{24404514895 23223104081
	26733218961 22901424206
	23194740344 24063270182
	26543255312}, {23223104081
	26733218961 22901424206
	23194740344 24063270182
	26543255312 23169771772},
	{26733218961 22901424206
	23194740344 24063270182
	26543255312 23169771772
	23788615880}, {22901424206
	23194740344 24063270182
	26543255312 23169771772
	23788615880 23522057990},
	{23194740344 24063270182
	26543255312 23169771772
	23788615880 23522057990
	22947868899}, {24063270182
	26543255312 23169771772
	23788615880 23522057990
	22947868899 23705631959},
	{26543255312 23169771772
	23788615880 23522057990
	22947868899 23705631959

Dokumen Uji	Dokumen Pemandang	
	24967182553},	{23169771772
	23788615880	23522057990
	22947868899	23705631959
	24967182553	27054500635},
	{23788615880	23522057990
	22947868899	23705631959
	24967182553	27054500635
	24077574926},	{23522057990
	22947868899	23705631959
	24967182553	27054500635
	24077574926	26700607505},
	{22947868899	23705631959
	24967182553	27054500635
	24077574926	26700607505
	24900645879},	{23705631959
	24967182553	27054500635
	24077574926	26700607505
	24900645879	23964649520},
	{24967182553	27054500635
	24077574926	26700607505
	24900645879	23964649520
	25458428033},	{27054500635
	24077574926	26700607505
	24900645879	23964649520
	25458428033	23026410141},
	{24077574926	26700607505
	24900645879	23964649520
	25458428033	23026410141
	24569585639},	{26700607505
	24900645879	23964649520

Dokumen Uji	Dokumen Pemanding
	25458428033 23026410141
	24569585639 25038882262}

H. Membentuk *Fingerprint*

Hasil proses pembentukan *fingerprint* ditampilkan pada Tabel 4.7. *Fingerprint* dipilih dari hasil nilai *hash* terkecil tiap *window*. Jika terdapat nilai *hash* terkecil yang sama pada *window* maka yang diambil untuk dijadikan *fingerprint* adalah *window* sebelah kanan.

Tabel 4. 7 Proses *Fingerprint*

Dokumen Uji	Dokumen Pemanding
23003449966 23220628354	23003449966 23220628354
23960433517 23224737912	23960433517 23224737909
23099117811 23011422306	23099117448 23011418331
23173678460 23994767148	23115471368 23102866356
22896958029 23145612397	23246079733 23122015013
23998872061 23848238210	24809393112 23136225748
23088327701 23011403116	23099116121 23011403711
	22901424206 22947868899
	23705631959 23964649520
	23026410141

I. Menghitung Kemiripan Menggunakan *Jaccard Coefficient*

Pada tahap ini berguna untuk mengetahui tingkat persentase kemiripan dokumen dari *fingerprint* yang dihasilkan pada metode *Winnowing*. Berikut merupakan perhitungan dari perhitungan dokumen uji dan dokumen pemanding di atas berdasarkan pada Persamaan 2.2 yang dituliskan pada Bab 2.

$$\text{Persentase Kemiripan} = \frac{3}{30} \times 100\% = 10\%$$

Berdasarkan perhitungan di atas, nilai 3 berasal dari jumlah *fingerprint* yang sama antara dokumen uji dan dokumen pemanding. *Fingerprint* yang sama yaitu 23003449966, 23220628354 dan 23960433517. Sedangkan nilai 30 merupakan total dari penggabungan *fingerprint*. *Fingerprint* yang digabungkan yaitu

23003449966, 23220628354, 23960433517, 23224737912, 23099117811, 23011422306, 23173678460, 23994767148, 22896958029, 23145612397 23998872061, 23848238210, 23088327701, 23011403116, 23224737909 23099117448, 23011418331, 23115471368, 23102866356, 23246079733 23122015013, 24809393112, 23136225748, 23099116121, 23011403711 22901424206, 22947868899, 23705631959, 23964649520 dan 23026410141. Jadi kesamaan antara dokumen uji dan dokumen pembanding menghasilkan persentase nilai 10%.

4.2 Analisis Pengguna

Pengguna aplikasi sistem pendeteksi plagiarisme dokumen teks lintas bahasa menggunakan metode *Winnowing* ini yaitu *user* dan *admin* yang masing-masing tugas, hak akses dan persyaratan khususnya dijabarkan pada Tabel berikut

Tabel 4. 8 Analisis Pengguna

Kategori	Tugas	Hak Akses	Persyaratan Khusus
Admin	Mengelola sistem secara keseluruhan	Semua fitur	<i>Username</i> dan <i>password</i> harus terdaftar di <i>database</i>
<i>User</i>	Melakukan perhitungan, input dokumen dan hapus dokumen maupun riwayat perhitungan	Fitur input dokumen teks, perhitungan tanpa dapat melihat tiap proses perhitungan dan fitur hapus	Pengguna umum

4.3 Analisis Kebutuhan Fungsional

Analisis kebutuhan fungsional dari sistem pendeteksi dokumen teks lintas bahasa menggunakan metode *Winnowing* adalah sebagai berikut:

- a. Data yang dibutuhkan adalah data dari jurnal yang terindikasi melakukan plagiasi pada jurnal lain maupun jurnal yang tidak terindikasi plagiasi
- b. Sistem dapat membaca dokumen berupa teks maupun PDF

- c. Sistem dapat menerjemahkan dokumen teks berbahasa Inggris ke bahasa Indonesia
- d. Sistem dapat membandingkan dokumen uji dengan dokumen pembanding
- e. Hasil kemiripan dokumen ditampilkan dalam bentuk persentase

4.4 Perancangan Detail Sistem

4.4.1 Perancangan Proses

Pada perancangan proses menggambarkan tentang lingkungan dari sistem penilaian plagiasi dokumen uji dengan dokumen pembanding. Perancangan proses sistem ini adalah sebagai berikut:

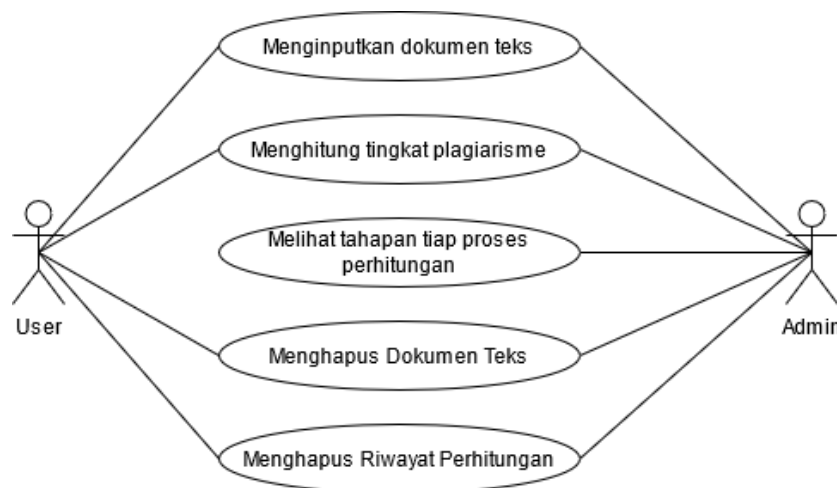
- a. *User* melakukan memasukkan data dokumen berupa teks atau *file* PDF jurnal
- b. Sistem akan menampilkan data dokumen yang sudah dimasukkan
- c. *User* dapat memilih dokumen uji dan dokumen pembanding untuk dilakukan perhitungan dokumen teks
- d. Sistem akan melakukan mendeteksi kemiripan dokumen menggunakan metode *Winnowing*
- e. Hasil akan ditampilkan berupa persentase kemiripan dokumen

Proses perancangan sistem perhitungan dokumen teks dari tahap input dokumen teks sampai hasil persentase kemiripan menggunakan *Jaccard Coefficient* digambarkan dalam bentuk diagram alir (*flowchart*) seperti pada Gambar 3.1. Proses kerja sistem secara umum yang terjadi yaitu memasukkan dokumen berupa teks ataupun *file* PDF, memilih dokumen uji dan pembanding, terjemahan dokumen bahasa Inggris ke bahasa Indonesia, *preprocessing* dokumen dengan *case folding*, *filtering* dan *stopword removal*, hitung kemiripan dokumen dengan *Winnowing* dan menghitung persentase kemiripan dengan *Jaccard Coefficient*.

4.4.2 Perancangan Use Case Diagram

Pada Gambar 4.1 berikut merupakan *use case diagram* untuk aplikasi sistem pendeteksi kemiripan dokumen teks. Di mana terdapat dua jenis pengguna yaitu *user* dan admin. Admin dapat menggunakan semua fitur seperti memasukkan dokumen teks, menghitung tingkat plagiarisme, melihat tahapan tiap proses perhitungan, menghapus dokumen teks dan menghapus riwayat perhitungan.

Sedangkan pada *user* dapat menggunakan fitur tersebut kecuali melihat tahapan tiap proses perhitungan. Sehingga *user* hanya dapat mengetahui hasil persentase kesamaan tiap dokumen yang diuji dan dibandingkan



Gambar 4. 1 *Use Case Diagram*

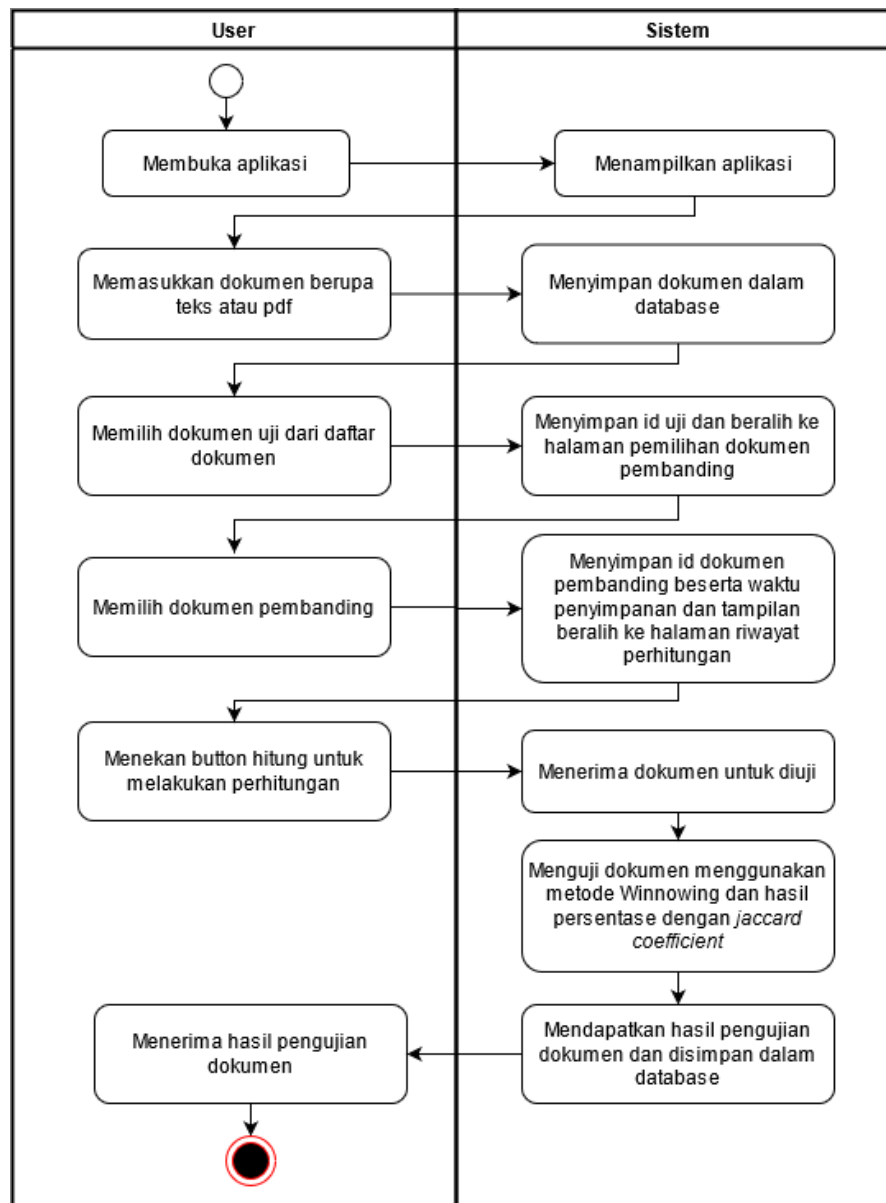
4.4.3 Perancangan *Activity Diagram*

Pada *activity diagram* menggambarkan aktivitas apa saja yang dilakukan oleh *user* dan sistem secara bertahap. Gambar 4.2 berikut merupakan *activity diagram* dari sistem pendeteksi plagiarisme dokumen teks lintas bahasa yang dibuat.

Aktivitas yang dilakukan pertama kali oleh *user* saat ingin menjalankan proses plagiarisme dokumen teks adalah membuka aplikasi. Lalu sistem akan menampilkan aplikasi pada *user* dan setelah aplikasi muncul pada tampilan *user* maka *user* dapat memasukkan dokumen berupa teks atau PDF. *User* dapat memasukkan 5 dokumen PDF sekaligus. Jika sudah selesai memasukkan dokumen, sistem akan menyimpan dokumen tersebut ke dalam *database*.

Dokumen yang berhasil disimpan oleh sistem di *database* akan ditampilkan pada *user* sehingga *user* dapat memilih dokumen mana yang akan dijadikan dokumen uji. Sistem akan menyimpan id dokumen uji dan tampilan halaman akan beralih ke halaman pemilihan dokumen pembanding. *User* dapat memilih dokumen pembanding dan jika pemilihan dokumen pembanding sudah selesai maka sistem akan menerima id dokumen pembanding yang nantinya akan digunakan untuk proses perhitungan. Saat proses perhitungan, sistem akan menguji dokumen menggunakan metode *Winnowing* dan dilanjutkan dengan perhitungan *Jaccard Coefficient* untuk memperoleh hasil persentase kemiripan. Jika proses perhitungan

sudah selesai maka sistem akan menyimpan hasil pengujian di *database* sehingga *user* dapat menerima hasil pengujian dokumen.

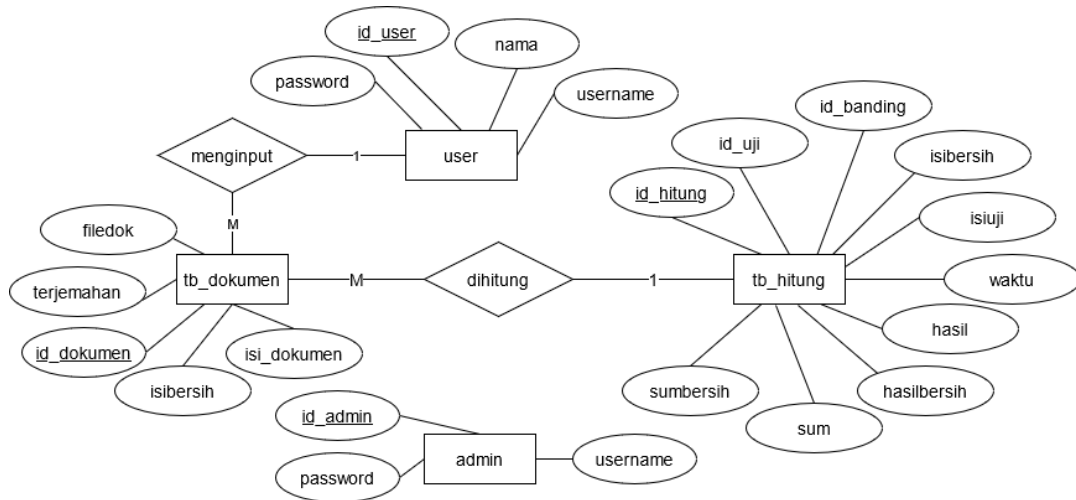


Gambar 4. 2 Activity Diagram

4.4.4 Perancangan Entity Relationship Diagram (ERD)

Pemodelan struktur data dan hubungan antar data pada sistem pendeteksi plagiarisme dokumen teks lintas bahasa ini digambarkan dengan *Entity Relationship Diagram (ERD)*. Pada Gambar 4.3 terdapat 4 entitas yaitu *tb_dokumen*, *tb_hitung*, *user* dan *admin*. Entitas *tb_dokumen* dihitung dan hasil perhitungannya disimpan pada entitas *tb_hitung* dan memiliki relasi 1 to M (*one to many*). Entitas *user* menginputkan dokumen dan akan disimpan pada entitas

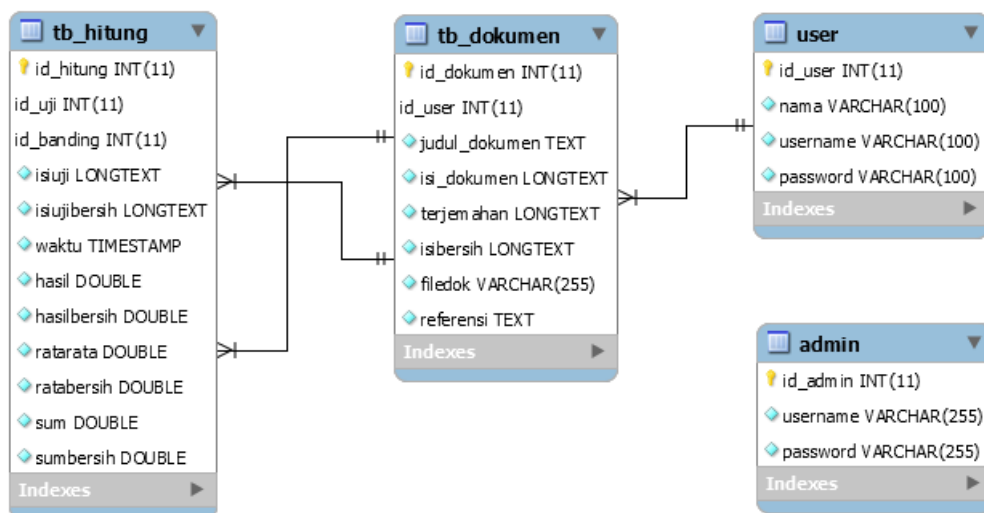
tb_dokumen dan memiliki relasi 1 to M (*one to many*). Sedangkan entitas admin tidak memiliki relasi dengan entitas lainnya.



Gambar 4. 3 Entity Relationship Diagram (ERD) Sistem

4.4.5 Perancangan Desain Database

Pada *database* yang dibangun terdapat relasi antara tabel **tb_hitung** dengan **tb_dokumen** dan **tb_dokumen** dengan **user**. Isi kolom *id_uji* dan *id_banding* pada tabel **tb_hitung** berdasarkan pada isi kolom *id_dokumen* pada tabel **tb_dokumen**. Kolom *id_user* pada tabel **tb_dokumen** berdasarkan pada isi *id_user* pada tabel **user**. Sedangkan pada tabel **admin** tidak memiliki relasi apapun dengan tabel lainnya. Berikut merupakan desain *database* yang digambarkan pada Gambar 4.4.



Gambar 4. 4 Perancangan Desain Database

4.4.6 Perancangan *Database*

Pada perancangan *database* memuat tentang penggunaan *database* pada sistem pendeteksi plagiarisme dokumen teks yang berguna untuk menyimpan data sehingga menghasilkan informasi. Pada sistem yang dibangun ini, terdapat 4 tabel yaitu admin, user, tb_dokumen dan tb_hitung. Berikut merupakan struktur dari masing-masing tabel.

A. Tabel admin

Pada tabel admin berguna untuk menyimpan daftar-daftar admin seperti *id_admin*, *username* dan *password*. Pada Tabel 4.9 berikut merupakan detail dari struktur tabel admin.

Keterangan : Berisi daftar-daftar admin

Primary key : *id_admin*

Tabel 4. 9 Tabel Admin

Nama Atribut	Tipe Data	Keterangan
<i>ID_ADMIN</i>	INT (11)	NOT NULL, PRIMARY_KEY
<i>USERNAME</i>	VARCHAR (255)	NOT NULL
<i>PASSWORD</i>	VARCHAR (255)	NOT NULL

B. Tabel user

Pada tabel user berguna untuk menyimpan daftar-daftar user seperti *id_user*, nama, *username* dan *password*. Berikut merupakan detail dari struktur tabel admin pada Tabel 4.10.

Keterangan : Berisi daftar-daftar user

Primary key : *id_user*

Tabel 4. 10 Tabel User

Nama Atribut	Tipe Data	Keterangan
<i>ID_USER</i>	INT (11)	NOT NULL, PRIMARY_KEY
<i>NAMA</i>	VARCHAR (100)	NOT NULL
<i>USERNAME</i>	VARCHAR (100)	NOT NULL
<i>PASSWORD</i>	VARCHAR (100)	NOT NULL

C. Tabel tb_dokumen

Pada tabel *tb_dokumen* berguna untuk menyimpan dokumen uji dan dokumen pembanding yang meliputi *id_dokumen*, *id_user*, *judul_dokumen*, *isi_dokumen*, *terjemahan*, *isibersih* dan *filedok*. Pada *isi_dokumen* merupakan isi keseluruhan dokumen, sedangkan pada *terjemahan* merupakan isi keseluruhan dokumen yang sudah diterjemahkan dan sudah dibatasi kalimat sebelum abstrak dan daftar pustaka dihilangkan. Pada *isibersih* berisi hasil isi terjemahan namun kalimat kutipan dan sitasi sudah dihilangkan. Pada Tabel 4.11 berikut merupakan detail struktur dari tabel *tb_dokumen*.

Keterangan : Berisi daftar dokumen uji dan dokumen pembanding

Primary key : *id_dokumen*

Tabel 4. 11 Tabel *tb_dokumen*

Nama Atribut	Tipe Data	Keterangan
ID_DOKUMEN	INT (11)	NOT NULL, PRIMARY KEY
ID_USER	INT (11)	NOT NULL
JUDUL_DOKUMEN	TEXT	NOT NULL
ISI_DOKUMEN	LONGTEXT	NOT NULL
TERJEMAHAN	LONGTEXT	NOT NULL
ISIBERSIH	TEXT	NOT NULL
FILEDOK	VARCHAR (255)	NOT NULL

D. Tabel *tb_hitung*

Pada tabel *tb_hitung* berisi hasil perhitungan dan id dokumen uji dan dokumen banding. Pada Tabel 4.12 berikut merupakan detail struktur tabel *tb_hitung*. Pada *hasil* dan *hasilbersih* merupakan hasil persentase dokumen uji dan dokumen pembanding menggunakan *jaccard coefficient*, sedangkan *sum* dan *sumbersih* berisi hasil jumlah dari *hasil* dan *hasilbersih* tiap dokumen. Setelah *hasil* *sum* dan *sumbersih* berhasil dilakukan maka dilanjutkan menghitung rata-rata dari *sum* dan *sumbersih*. Hasil akhir kesamaan dokumen ditunjukkan ada kolom *ratarata* dan *ratabersih*

Keterangan : Berisi riwayat perhitungan

Primary key : *id_hitung*

Tabel 4. 12 Tabel tb_hitung

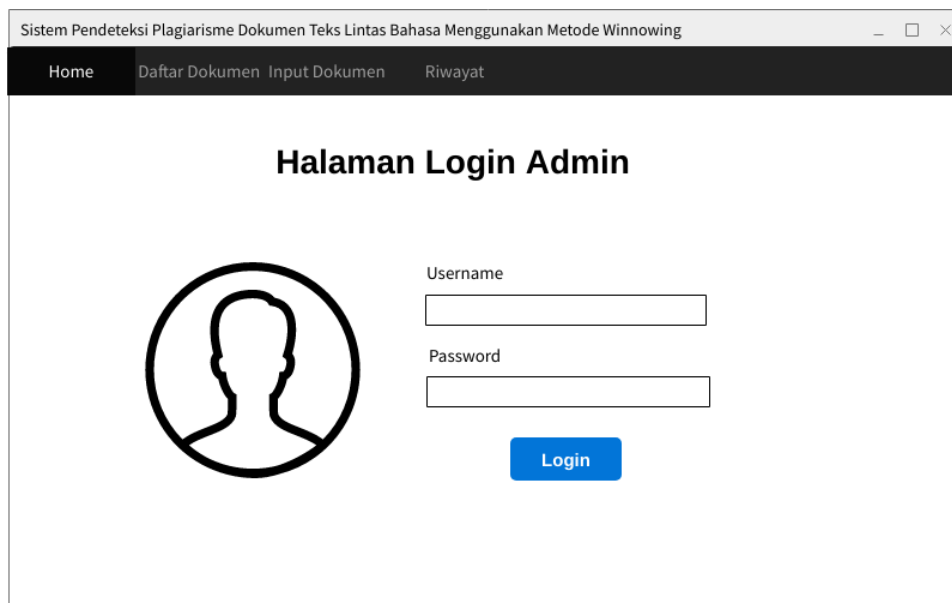
Nama Atribut	Tipe Data	Keterangan
ID_HITUNG	INT (11)	NOT NULL, PRIMARY KEY
ID_UJI	INT (11)	NOT NULL
ID_BANDING	INT (11)	NOT NULL
ISIUI	LONGTEXT	NOT NULL
ISIUIBERSIH	LONGTEXT	NOT NULL
WAKTU	TIMESTAMP	NOT NULL
HASIL	DOUBLE	NOT NULL
HASILBERSIH	DOUBLE	NOT NULL
SUM	DOUBLE	NOT NULL
SUMBERSIH	DOUBLE	NOT NULL
RATARATA	DOUBLE	
RATABERSIH	DOUBLE	

4.4.7 Perancangan Antarmuka Pengguna

Pada perancangan antarmuka aplikasi yang dibuat ini terdiri dari halaman admin dan halaman *user*. Halaman admin dapat diakses apabila *username* dan *password* sudah terdaftar pada *database*.

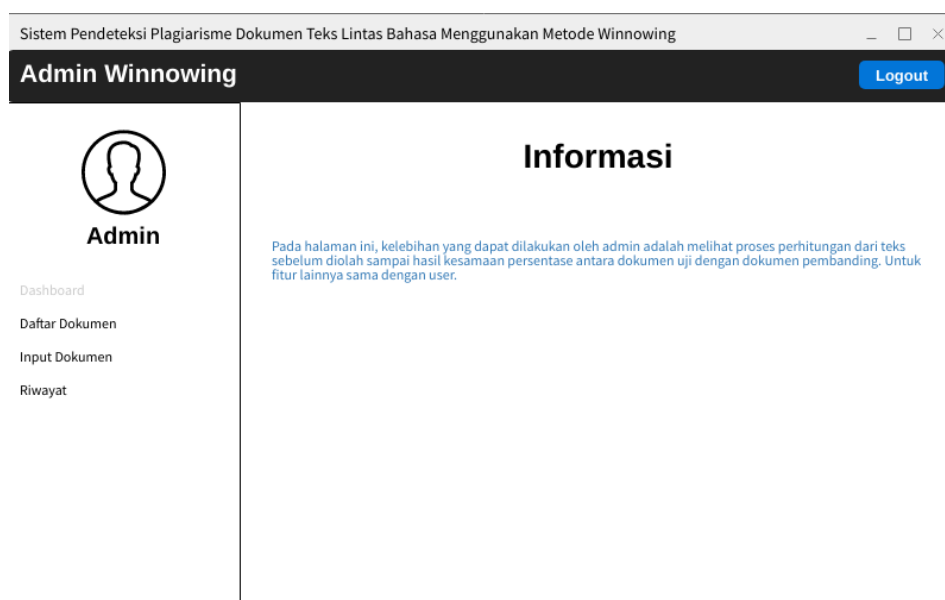
A. Halaman Beranda Utama atau *Home*

Pada halaman beranda utama *user* ditampilkan pada Gambar 4.5 berikut di mana terdapat menu pada navbar yaitu *home*, dokumen referensi, daftar dokumen, riwayat dan login admin. Halaman beranda utama atau *home* adalah halaman yang muncul pertama kali saat pengguna membuka aplikasi sistem pendeteksi plagiarisme. Pengguna dapat login terlebih dahulu sebelum dapat melakukan perhitungan. Untuk pengguna yang belum mempunyai akun dapat melakukan register terlebih dahulu. Pengguna yang tidak melakukan login hanya dapat melihat daftar dokumen referensi, daftar dokumen dan riwayat perhitungan.



Gambar 4. 5 Antarmuka Halaman Beranda Utama *User*

Sedangkan halaman *home* yang ditampilkan pada admin ditampilkan seperti pada Gambar 4.6. Pada halaman *home* admin terdapat *button logout*, *sidebar* yang berisi menu *dashboard*, daftar dokumen, input dokumen dan riwayat. Halaman *home* admin juga menampilkan informasi dari aplikasi yang dibuat beserta tahapan perhitungan *WInnowing*.



Gambar 4. 6 Antarmuka Halaman Beranda Utama Admin

B. Halaman Daftar Dokumen

Pada Gambar 4.7 berikut merupakan tampilan antarmuka halaman daftar dokumen pada *user*. Pada halaman ini pengguna dapat melihat dokumen apa saja

yang pernah diinputkan. Pengguna dapat melakukan menghapus pada dokumen yang diinginkan dan melihat perbandingan teks keseluruhan dengan teks yang diterjemahkan dan batasan teks tanpa daftar pustaka.

Nomor	Judul Dokumen	Isi Dokumen	Terjemahan	Aksi
1	50-Article Text-98-1-10-...	IMPLEMENTASI TOKENIZING PL...	syarat lulus bagi program sarja...	Hitung Hapus
2	243-Article Text-521-1-...	Abstrak Peraturan Menteri Pend...	peraturan menteri pendidikan ...	Hitung Hapus
3	69-1-496-1-10-20181008	ABSTRACT The search for simila...	pencarian kesamaan judul tuga...	Hitung Hapus

Gambar 4. 7 Antarmuka Halaman Daftar Dokumen *User*

Sedangkan tampilan halaman daftar dokumen pada admin ditampilkan seperti pada Gambar 4.8. Tampilan pada halaman admin tidak jauh berbeda seperti yang ditampilkan pada *user* yaitu menampilkan judul dokumen, isi dokumen, isi terjemahan beserta button aksi hitung dan hapus.

No	Judul Dokumen	Isi Dokumen	Terjemahan	Aksi
1	50-Article Text-9-...	IMPLEMENTASI TOKENIZI...	syarat lulus bagi program ...	Hitung Hapus
2	243-Article Text-...	Abstrak Peraturan Menteri ...	peraturan menteri pendidi...	Hitung Hapus
3	69-1-496-1-10-2-...	ABSTRACT The search for s...	pencarian kesamaan judul...	Hitung Hapus

Gambar 4. 8 Antarmuka Halaman Daftar Dokumen Admin

C. Halaman Pilih Dokumen Pemanding

Tampilan desain antarmuka halaman untuk memilih dokumen pemanding pada *user* seperti pada Gambar 4.9. *User* dan admin dapat memilih dokumen pemanding dengan menekan pada *checkbox*.

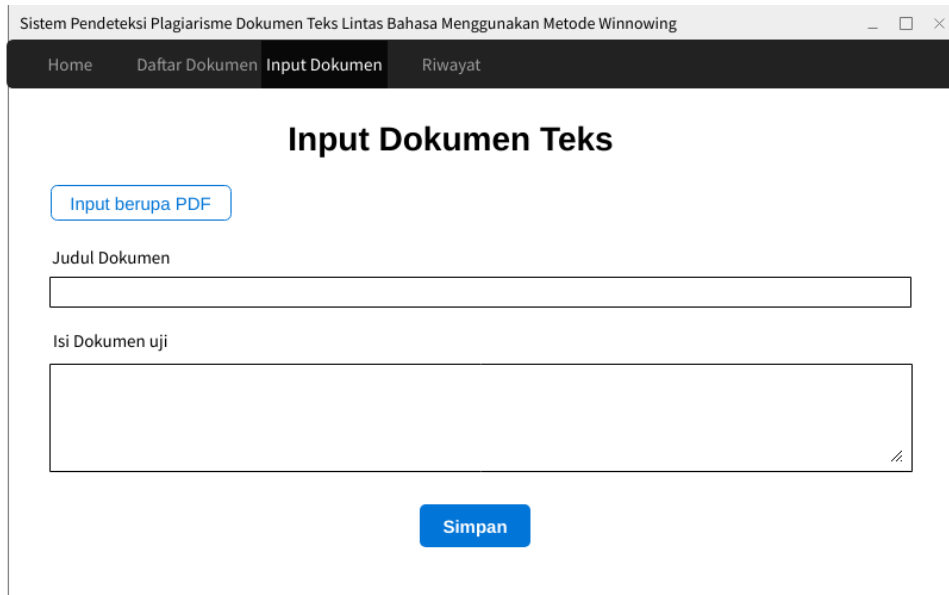
Gambar 4. 9 Antarmuka Halaman Pilih Dokumen Pemanding *User*

Sedangkan desain antarmuka tampilan halaman pilih dokumen pemanding pada admin dapat dilihat pada Gambar 4.10 berikut.

Gambar 4. 10 Antarmuka Halaman Pilih Dokumen Pemanding Admin

D. Halaman Input Dokumen

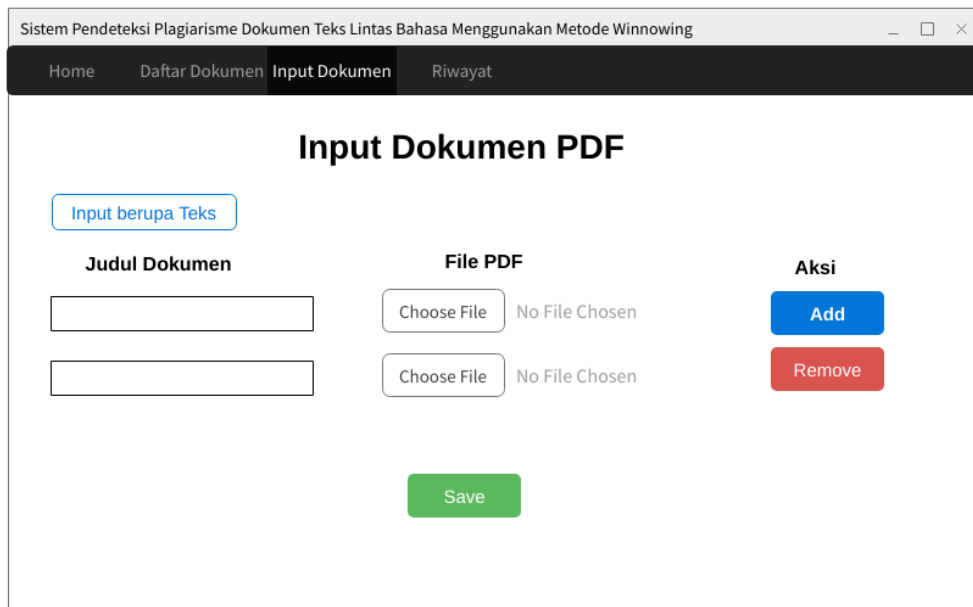
Pengguna dapat menginputkan dokumen berupa teks maupun *file* PDF pada halaman input dokumen. Jika *user* menginputkan dokumen berupa teks maka tampilan desain antarmuka seperti pada Gambar 4.11 berikut.



The screenshot shows a web browser window titled "Sistem Pendeteksi Plagiarisme Dokumen Teks Lintas Bahasa Menggunakan Metode Winnowing". The navigation bar includes "Home", "Daftar Dokumen", "Input Dokumen", and "Riwayat". The main heading is "Input Dokumen Teks". There is a button labeled "Input berupa PDF". Below it, there are two input fields: "Judul Dokumen" and "Isi Dokumen uji". At the bottom, there is a blue "Simpan" button.

Gambar 4. 11 Antarmuka Halaman Input Dokumen Teks *User*

Jika *user* ingin menginputkan dokumen berupa *file* PDF maka tampilan desain antarmuka seperti Gambar 4.12 berikut.



The screenshot shows a web browser window titled "Sistem Pendeteksi Plagiarisme Dokumen Teks Lintas Bahasa Menggunakan Metode Winnowing". The navigation bar includes "Home", "Daftar Dokumen", "Input Dokumen", and "Riwayat". The main heading is "Input Dokumen PDF". There is a button labeled "Input berupa Teks". Below it, there are two input fields for "Judul Dokumen". To the right, there is a "File PDF" section with two "Choose File" buttons and "No File Chosen" text. To the right of that, there is an "Aksi" section with "Add" and "Remove" buttons. At the bottom, there is a green "Save" button.

Gambar 4. 12 Antarmuka Halaman Input Dokumen PDF *User*

Sedangkan tampilan desain antarmuka halaman untuk memasukkan dokumen berupa teks pada admin ditampilkan seperti pada Gambar 4.13 berikut.

Gambar 4. 13 Antarmuka Halaman Input Dokumen Teks Admin

Jika admin memasukkan dokumen berupa PDF maka tampilan desain antarmuka seperti pada Gambar 4.14 berikut.

Gambar 4. 14 Antarmuka Halaman Input Dokumen Teks PDF Admin

E. Halaman Riwayat

Pada desain antarmuka halaman riwayat yang ditampilkan pada *user* dapat dilihat seperti Gambar 4.15 yang menampilkan hasil riwayat perhitungan yang

dilakukan. Untuk mengetahui kesamaan dokumen uji dan pembanding keseluruhan, pengguna dapat menekan *button* Hitung. Jika pengguna ingin mengetahui kesamaan dokumen uji dan pembanding tanpa kutipan dan sitasi, maka pengguna dapat menekan *button* Hitung Bersih.

No	Waktu	ID Dokumen Uji	Jumlah Dokumen Banding	Hasil	Hasil Bersih	Aksi
1	2021-04-27 02:26:27	1	3	31%	29%	Hitung Hitung Hapus
2	2021-04-26 10:17:31	2	4	20.4%	20.5%	Hitung Hitung Hapus
3	2021-04-23 12:31:13	3	2	23.6%	24%	Hitung Hitung Hapus

Gambar 4. 15 Antarmuka Halaman Riwayat *User*

Sedangkan desain antarmuka halaman riwayat yang ditampilkan pada admin seperti pada Gambar 4.16 berikut.

No	Waktu	ID Dokumen Uji	Jumlah Dokumen...	Hasil	Hasil Bersih	Aksi
1	2021-04-27 02:...	1	3	31%	29%	Hitung Hitung Hapus
2	2021-04-26 10:...	2	4	20.4%	20.5%	Hitung Hitung Hapus
3	2021-04-23 12:...	3	2	23.6%	24%	Hitung Hitung Hapus

Gambar 4. 16 Antarmuka Halaman Riwayat Admin

F. Halaman Detail Hasil Perhitungan

Pada halaman ini menampilkan tahapan pemrosesan dokumen teks dari teks sebelum diolah sampai hasil persentase yang dihasilkan dari *Jaccard Coefficient*. Proses yang ditampilkan pada halaman *user* lebih sederhana dibandingkan yang ditampilkan pada halaman admin. Berikut merupakan tampilan pada halaman *user* seperti pada Gambar 4.17.

Persentase Kemiripan Jaccard Coefficient			
No	ID Uji	ID Banding	Hasil Persentase
1	421	419	6.3%
2	421	420	5.8%

Hasil Akhir			
No	Jumlah Dokumen Pembanding	Hasil Sum	Hasil Rata-Rata
1	2	12.1%	6.1%

[Kembali](#)

Gambar 4. 17 Antarmuka Halaman Detail Hasil Perhitungan *User*

Sedangkan desain antarmuka halaman detail hasil perhitungan pada admin seperti Gambar 4.18 berikut.

Teks belum diolah				
No	ID Uji	ID Banding	Teks Uji	Teks Banding
1	421	419	Sistem esai pada penelitian ini dil... Algoritma winnowing adalah algo...	
2	421	420	Sistem esai pada penelitian ini dil... Maraknya kasus penjiplakan yang...	

Preprocessing (Casefolding)				
No	ID Uji	ID Banding	Teks Uji	Teks Banding
1	421	419	sistem esai pada penelitian ini dil... sgoritma winnowing adalah algo...	
2	421	420	sistem esai pada penelitian ini dil... maraknya kasus penjiplakan yan...	

Gambar 4. 18 Antarmuka Halaman Detail Hasil Perhitungan Admin