



BAB IV

ANALISA DAN PERANCANGAN

Analisa dan perancangan adalah tahapan penting dalam melakukan sebuah penelitian. Tahap ini dilakukan dalam penelitian untuk memahami kebutuhan yang dibutuhkan untuk membangun aplikasi. Tahapan ini merupakan proses penyesuaian masalah yang ada dalam penelitian.

4.1 Analisa Kebutuhan Data

Data yang digunakan pada penelitian merupakan data Tugas Akhir mahasiswa Jurusan Teknik Informatika dari tahun 2016 sampai dengan awal 2020 dengan total jumlah data awal sebanyak 412 *record* yang didapat dari *repository* UIN SUSKA RIAU. Atribut yang didapat yaitu nama mahasiswa, judul Tugas Akhir, bidang keahlian. Atribut ini dijelaskan pada tabel 4.1 berikut :

Tabel 4.1 Tabel Data Atribut

No	Atribut	Keterangan
1	Nama Mahasiswa	Nama Lengkap Mahasiswa
2	Judul Tugas Akhir	Judul Tugas Akhir mahasiswa
3	Bidang Keahlian	Pembagian Bidang keahlian Tugas Akhir mahasiswa

Berdasarkan penjelasan diatas, keseluruhan data yang digunakan pada penelitian ini dijelaskan pada tabel 4.2 berikut :



Tabel 4.2 Tabel Bidang Keahlian

No	Bidang Keahlian	Kode
1	Computer Science (CS)	1
2	Technology Information (TI)	2

Tabel 4.3 data Judul Tugas Akhir

NO	penulis	Judul	Kel as	tahu n
1	Anissah Putri	Simulasi Hitung Warisan Berbasis Multimedia Berdasarkan Hukum Islam (Fara'id) Menggunakan Aturan (Rule Based)	TI	2018
2	Irni Mayriantika	Perancangan Manajemen Layanan Teknologi Informasi Menggunakan Framework Information Technology Infrastructure Library (ITIL) V3 (Studi Kasus : Dinas Komunikasi Informatika dan Statistika Provinsi Riau)	TI	2018
3	Afri Rahmad	Perancangan Data Center Menggunakan Framework TIA-942 Pada Sekretariat Daerah (SETDA) Kabupaten Siak	TI	2018
4	Satria Nugraha	APLIKASI PEMANTAUAN KEGIATAN BAKAL CALON ANGGOTA DEWAN BERBASIS ANDROID (STUDI KASUS: DPW PKS RIAU)	TI	2018
5	Aby Wahyu	Aplikasi Real Count (Studi Kasus: PILGUB Riau 2018)	TI	2018
6	Dhea Karina Rusdi	RANCANG BANGUN SISTEM INFORMASI MANAJEMEN PROYEK MENGGUNAKAN PROJECT MANAGEMENT BODY OF KNOWLEDGE (PMBOK) 5	TI	2018



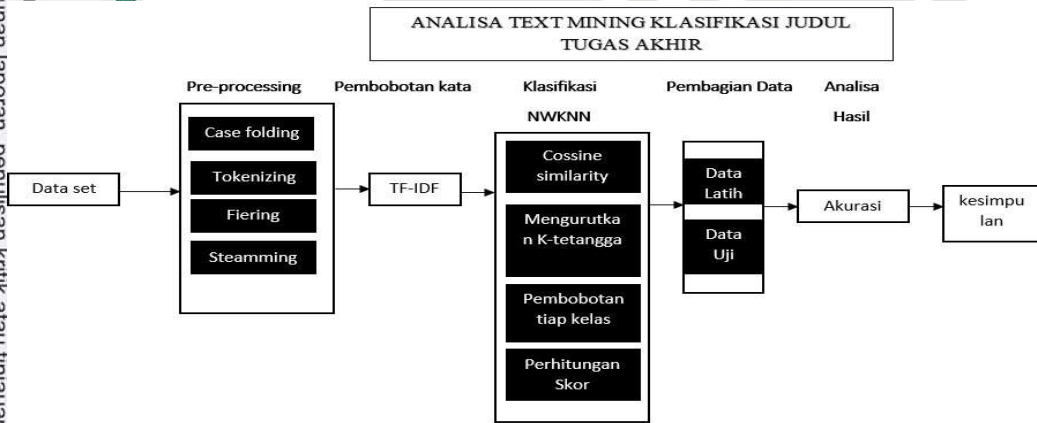
<p>7</p> <p>1. Ditarang-mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber.</p> <p>a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.</p> <p>b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.</p> <p>2. Dilarang-mengumumkan dan memperdayak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.</p>	<p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p> <p>© Hak Cipta Dilindungi Undang-Undang</p>	<p>DAYU ANGGARA SAPUTRA</p> <p>Aggy Muhammad Nuhammad</p> <p>Kuni Alya Faqriya</p> <p>ANDRI GUNAWAN</p> <p>Rizky Yananda</p> <p>RISWANTO</p> <p>TRI SIDDIK MUHAMMAD</p> <p>ODHI AHMAD HIDAYAT</p> <p>SAFRIDA IKA GUSELIANTO</p> <p>BOBBY KURNIAWAN</p> <p>FIKR SAPUTRA</p>	<p>RANCANG BANGUN APLIKASI PENGENALAN DAN PENCARIAN GEDUNG KAMPUS UIN SUSKA RIAU MENGGUNAKAN TEKNOLOGI VIRTUAL REALITY (VR) 360 DAN ALGORITMA A-STAR BERBASIS ANDROID</p> <p>APLIKASI AL-QUR'AN DIGITAL UNTUK MEMBANTU DALAM MENGULANG HAFALAN AL-QUR'AN MENGGUNAKAN GOOGLE SPEECH API</p> <p>Aplikasi Guide Hajj & Umrah pada Perangkat Android dengan Menggunakan Google Speech API</p> <p>APLIKASI M-SHALAT SEBAGAI TUNTUNAN PELAKSANAAN IBADAH SHALAT BERBASIS ANDROID</p> <p>APLIKASI DO'A SEHARI-HARI BERBASIS ANDROID DENGAN GOOGLE SPEECH API</p> <p>APLIKASI LAYANAN SMART-LIBRARY PERPUSTAKAAN (STUDI KASUS DINAS PERPUSTAKAAN DAN KEARSIPAN KOTA PEKANBARU)</p> <p>Aplikasi Market-Rental Kendaraan</p> <p>RANCANG BANGUN APLIKASI TEMU APP MENGGUNAKAN PLATFORM ANDROID</p> <p>APLIKASI PENDISTRIBUSIAN BENIH TANAMAN DIFASILITASI LOCATION BASED SERVICE (LBS) DENGAN GEOCODING MAPS</p> <p>RANCANG BANGUN SMART POSYANDU DENGAN MENGGUNAKAN LOCATION BASED SERVICE (LBS) BERBASIS ANDROID</p> <p>RANCANG BANGUN APLIKASI GAME EDUKASI BELAJAR MEMBACA DAN MENGHAFAKAL DO'A AGAMA ISLAM UNTUK ANAK-ANAK</p>	<p>TI</p> <p>TI</p> <p>TI</p> <p>TI</p> <p>TI</p> <p>TI</p> <p>TI</p> <p>TI</p> <p>TI</p> <p>TI</p> <p>TI</p> <p>TI</p> <p>TI</p> <p>TI</p>	<p>2018</p> <p>2018</p> <p>2018</p> <p>2018</p> <p>2018</p> <p>2018</p> <p>2018</p> <p>2018</p> <p>2018</p> <p>2018</p> <p>2018</p> <p>2018</p> <p>2018</p> <p>2018</p>
--	---	--	--	---	---



18. HESTI JUMADHA RAHMI 19. MACHFUD LAH 20. VERRY DWI SETIAWAN	Aplikasi Tahfizh Al-Qur'an dengan Koreksi dan Evaluasi Berbasis Android	TI	2018
	RANCANG BANGUN APLIKASI PELAYANAN KELURAHAN (E-VILLAGE) BERBASIS WEB STUDI KASUS : KELURAHAN TUAH MADANI	TI	2018
	IMPLEMENTASI PERATURAN MENTERI PENDAYAGUNAAN APARATUR NEGARA DAN REFORMASI BIROKRASI NOMOR 14 TAHUN 2017 PADA SISTEM PENILAIAN TINGKAT KEPUASAN PASIEN RUMAH SAKIT UMUM DAERAH ARIFIN ACHMAD	CS	2018

Analisis Text Mining Klasifikasi Data Tugas Akhir

Text mining mempunyai definisi sebagai menambang data berupa teks dimana data biasanya di dapat dari suatu dokumen dengan tujuan mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen. Secara umum sistem ini dibagi menjadi beberapa tahapan proses, pada gambar dibawah ini merupakan blok diagram proses kerja pada text mining klasifikasi Data Tugas Akhir.

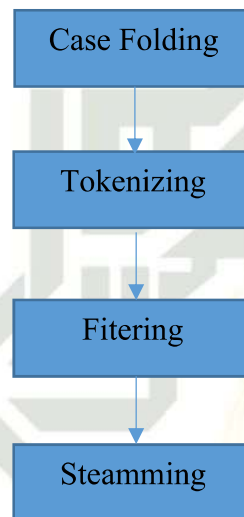


Gambar 4.1 Proses Kerja Text Mining Klasifikasi Judul Tugas Akhir



4.3 Pre-Processing

Tahapan *Pre-Processing* dalam *Text Mining* bertujuan untuk mempersiapkan data sebelum diproses pada langkah selanjutnya. Pada tahapan ini juga merupakan tahapan untuk membersihkan simbol dan tanda baca, mengganti kata baku menjadi tidak baku serta mengubah semua huruf ke dalam huruf kecil. Urutan tahapan *pre-processing* dapat dilihat pada gambar berikut :



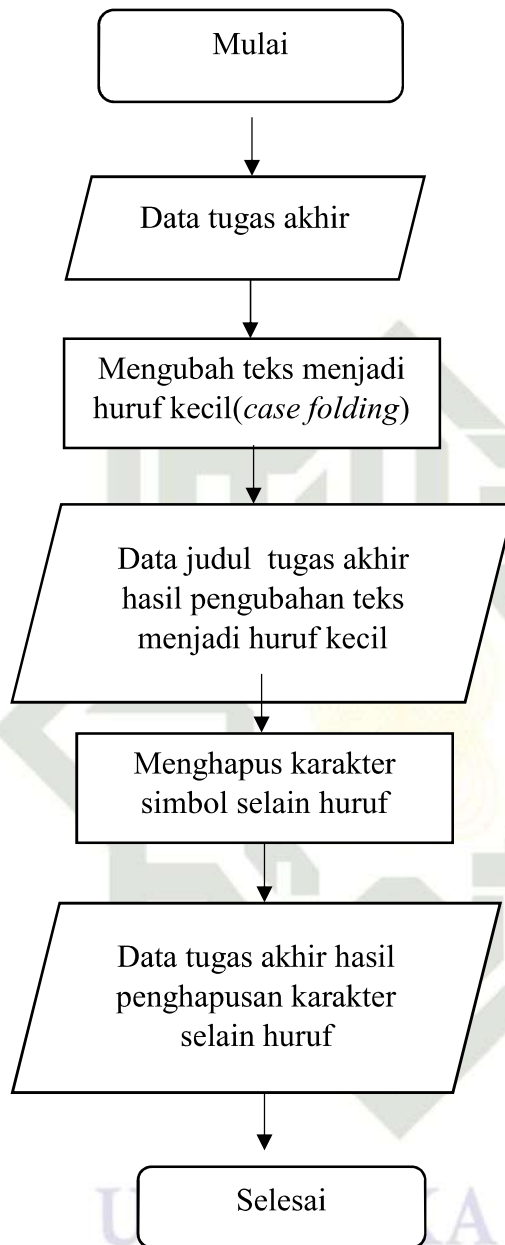
Gambar 4.2 Tahapan Pre-processing

4.3.1 Case folding

Tidak semua dokumen teks konsisten dalam penggunaan huruf kapital. Oleh karena itu, peran Case Folding dibutuhkan dalam mengkonversi keseluruhan teks dalam dokumen menjadi suatu bentuk standar (biasanya huruf kecil atau lowercase). Sebagai contoh, user yang ingin mendapatkan informasi KOMPUTER dan mengetik KOMPOTER, KomPUter atau komputer, tetap diberikan hasil retrieval yang sama yakni komputer. Case folding adalah mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf 'a' sampai dengan 'z' yang diterima. Karakter selain huruf dihilangkan dan dianggap delimiter. Berikut adalah flowchart tahapan dari *case folding* :

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 4.3 Tahapan Case Folding

Pada tahapan ini, ada beberapa aturan proses agar hasil *case folding* dapat sesuai dengan yang diharapkan, adapun aturan tersebut sebagai berikut :



Tabel 4.4 Aturan Case Folding

Kondisi	Aksi
Inputan data judul tugas akhir memiliki huruf kapital [A.....Z]	maka akan mengubah semua inputan tersebut menjadi huruf kecil [a.....z]
Inputan data judul tugas akhir memiliki karakter simbol	Maka akan menghapus karakter simbol tersebut dari inputan
Inputan data judul tugas akhir memiliki huruf kecil	tidak ada aksi
inputan data judul tugas akhir memiliki spasi	Tidak ada aksi

Karakter-karakter simbol yang akan dihapus atau dianggap sebagai pemisah kata, dapat dilihat pada tabel berikut :

Tabel 4.5 Karakter Simbol

Karakter simbol		
^	,	:
@	(;
%)	
\$	-	
#	<	}
'	>	{
~	+	\
&	=	/
.	!	

berikut ini contoh tahapan *case folding* yang akan diilustrasikan pada tabel dibawah ini :

Tabel 4.6 Case Folding

CONTOH DATA					CASE FOLDING				
No	Judul	Kelas	Kode	Dokumen	No	Judul	Kelas	Kode	Dokumen
	KLASIFIKASI SENTIMEN TERHADAP BUKALAPAK DENGAN MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER	CS	1	D3	3	klasifikasi sentimen terhadap bukalapak dengan menggunakan metode naïve bayes classifier	CS	1	D3

Sample Data 2018-2019

Hak Cipta Dilindungi Undang-Undang
 a. Pengutipan hanya untuk keperluan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 Dilarang mengutip dan/atau menyalin sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak cipta milik UIN

University of Sultan Syarif Kasim Riau



Hak Cipta Dilindungi Undang-undang

yang

5

5

5

5

5

5

APLIKASI REAL COUNT (STUDI KASUS: PILGUB RIAU 2018) COMPONENT ANALYSIS (PCA)

OPTIMASI FUZZY TIME SERIES MENGGUNAKAN PARTICLE SWARM OPTIMIZATION UNTUK MEMPREDIKSI KUALITAS UDARA DI KOTA PEKANBARU

TI	2	D4	4	aplikasi real count (studi kasus: pilgub riau 2018)	TI	2	D4
CS	1	D5	5	optimasi fuzzy time series menggunakan an particle swarm optimizatio n untuk mempredik si kualitas udara di kota pekanbaru	CS	1	D5

1. Ditaring-mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



KLASIFIKASI EMOSI PADA TWEET MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

1. Ditaring mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

klasifikasi emosi pada tweet menggunakan metode support vector machine	?	?	D6	6	?	?	D6
--	---	---	----	---	---	---	----



IMPLEMENTASI PERATURAN MENTERI KEMENDAYAGUNAAN APARATUR NEGARA DAN REFORMASI BIROKRASI NOMOR 14 TAHUN 2017 PADA SISTEM PENILAIAN TINGKAT KEPUASAN PASIEN RUMAH SAKIT UMUM DAERAH ARIFIN ACHMAD

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengemukakan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

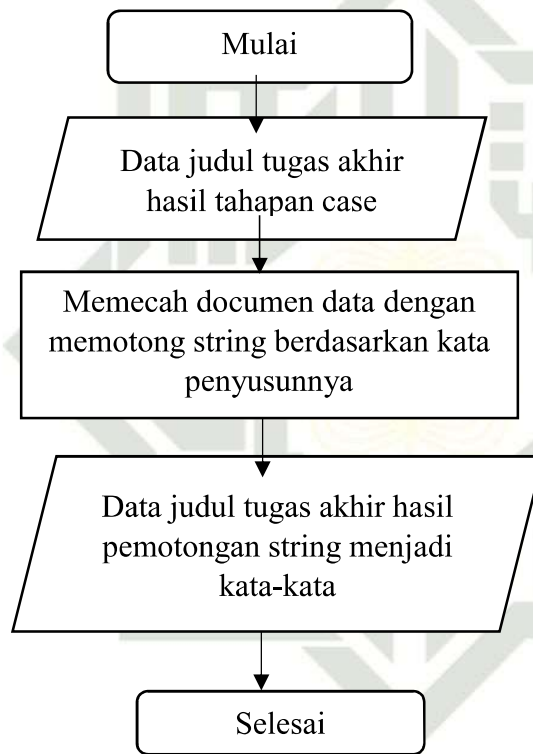
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

CS	1	D7	7	implementasi peraturan menteri pendayagunaan aparatur negara dan reformasi birokrasi nomor 14 tahun 2017 pada sistem penilaian tingkat kepuasan pasien rumah sakit umum daerah arifin achmad	CS	1	D7
----	---	----	---	--	----	---	----



4.3.2 Tokenizing

Tahapan *tokenizing* merupakan tahap pemotongan *string* input berdasarkan tiap kata yang menyusunnya. *Tokenizing* berdasarkan garis besar memecah sekumpulan karakter atau teks kedalam suatu kata. Contoh dari karakter seperti enter, tabulasi, spasi dan sebagainya sebagai pemisah kata. Kemudian untuk titik dua (:), petik tunggal (‘), semikolon(;) dan sebagainya, dapat memiliki peran yang cukup banyak sebagai pemisah kata. Proses *tokenizing* ini mempunyai alur yang digambarkan pada flowchart berikut ini :



Gambar 4.4 flowchart tahapan tokenizing

Pada tahapan ini dilakukan pemecahan de Tugas Akhir dari data judul tugas akhir menjadi bab-bab, para-grap, kalimat, dan menjadi kata-kata dengan memotong string dari

penyusunnya. Ada beberapa aturan proses agar hasilnya sesuai dengan yang diinginkan.

Ada beberapa aturan tersebut sebagai berikut:

Aturan proses tokenizing

Kondisi	Aksi
Jika inputan data judul tugas akhir bertitik spasial	Maka akan memecah dari deTugas Akhir data judul tugas akhir menjadi bab-bab per bagian kata atau string
Jika inputan data judul tugas akhir memiliki huruf	Tidak ada aksi

Operasian tahapan *Tokenizing* pada data judul Tugas Akhir dapat dilihat pada table dibawah

Table 4.8 Tabel Tokenizing

Judul tugas akhir	D3	D4	D5	D6	D7
	klasifikasi sentimen terhadap bukalapak dengan menggunakan metode naïve bayes classifier	aplikasi real count (studi kasus: pilgub riau 2018) component analysis (pca)	optimasi fuzzy time series menggunakan n particle swarm optimization untuk memprediksi i kualitas	klasifikasi emosi pada tweet menggunakan metode support vector machine	implementasi peraturan menteri pendayagunaan aparatur negara dan reformasi birokrasi nomor 14 tahun 2017



<p>© Hak cipta milik UIN Suska</p> <p>Hak Cipta Dilindungi Undang-Undang</p> <ol style="list-style-type: none"> 1. Dilarang mengutip sebagian atau seluruh a. Pengutipan hanya untuk kepentingan pribadi b. Pengutipan tidak merugikan kepentingan umum <p>2. Dilarang menggunakan dan memperbanyak atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.</p>	<p>udara di kota pekanbaru</p>			<p>pada sistem penilaian tingkat kepuasan pasien rumah sakit umum daerah arifin achmad</p>
<p>TOKENIZI NG</p>	<p>TOKENIZING</p>	<p>TOKENIZI NG</p>	<p>TOKENIZI NG</p>	<p>TOKENIZI NG</p>
<p>[klasifikasi] [sentimen] [terhadap] [bukalapak] [metode] [naive] [bayes] [classifier]</p>	<p>[Aplikasi][real] [count] [studi] [kasus] [pilgub][riau] [2018][component][an alysis] [pca]</p>	<p>[optimasi] [fuzzy] [time] [series] [menggunakan] [particle] [swarm] [optimizatio n] [untuk] [mempredik si] [kualitas] [udara] [di] [kota] [pekanbaru]</p>	<p>[klasifikasi] [emosi][pada] [tweet] [menggunakan] [metode support] [vector machine]</p>	<p>[implementas i][peraturan] [menteri] [pendayaguna an][aparatur] [negara] [dan] [reformasi] [birokrasi] [nomor] [14] [tahun][2017][pada] [system] [penilaian] [tingkat] [[kepuasan]</p>

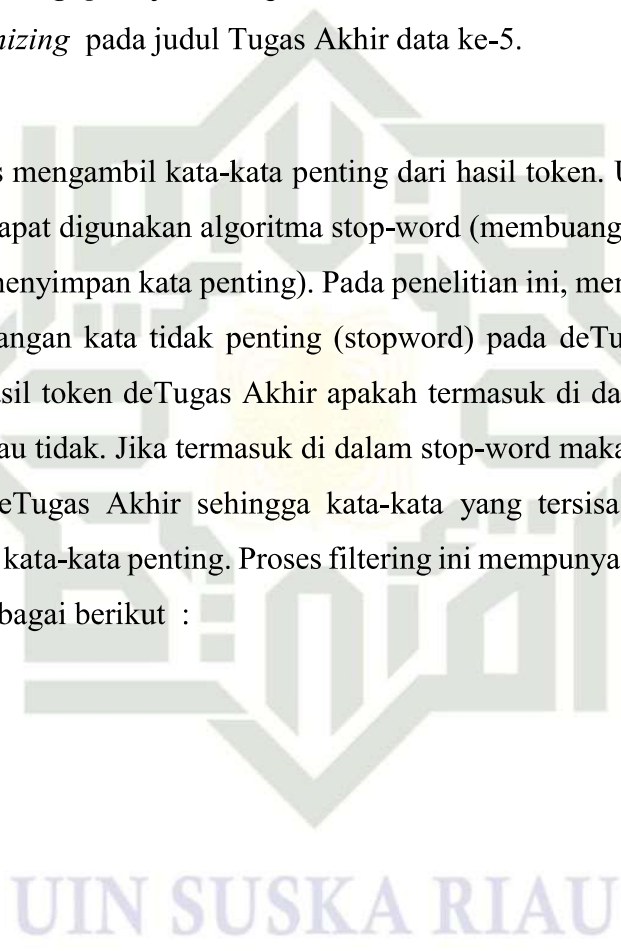


Hak Cipta Dilindungi Undang-Undang
 1. Dilarang mengutip sebagian atau seluruhnya dan untuk tujuan komersial atau penyalinan untuk umum tanpa mencantumkan dan menyebutkan sumber.
 2. Pengutipan hanya untuk keperluan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 3. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 4. Dilarang mengizinkan kepada pihak ketiga untuk menjual atau menyewakan data dan bentuk apapun tanpa izin UIN Suska Riau.

- Keterangan :
- D₁ merupakan hasil *tokenizing* pada judul Tugas Akhir data ke-1
 - D₂ merupakan hasil *tokenizing* pada judul Tugas Akhir data ke-2
 - D₃ merupakan hasil *tokenizing* pada judul Tugas Akhir data ke-3
 - D₄ merupakan hasil *tokenizing* pada judul Tugas Akhir data ke-4
 - D₅ merupakan hasil *tokenizing* pada judul Tugas Akhir data ke-5.

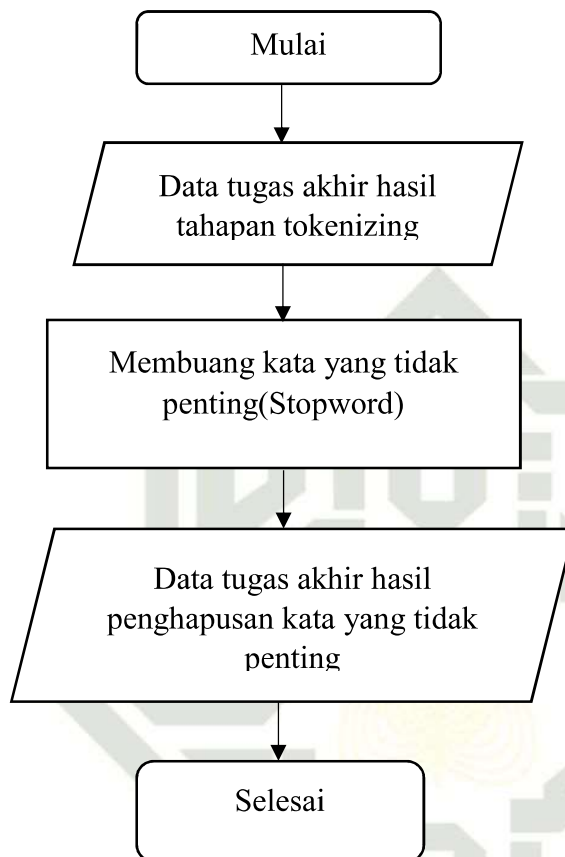
4.3 Filtering

Filtering merupakan proses mengambil kata-kata penting dari hasil token. Untuk bisa melakukan analisis filtering ini dapat digunakan algoritma stop-word (membuang kata yang kurang penting) atau word-list (menyimpan kata penting). Pada penelitian ini, menggunakan metode stop-word yaitu penghilangan kata tidak penting (stopword) pada deTugas Akhir atau pengecekan kata-kata hasil token deTugas Akhir apakah termasuk di dalam daftar kata tidak penting (stop-word) atau tidak. Jika termasuk di dalam stop-word maka kata-kata tersebut akan di-remove dari deTugas Akhir sehingga kata-kata yang tersisa di dalam deTugas Akhir di anggap sebagai kata-kata penting. Proses filtering ini mempunyai alur yang digambarkan pada gambar 4.5 sebagai berikut :



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber.
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 4.5 flowchart tahapan filtering

Pada tahap filtering ini menggunakan library sasatrawi untuk pembentukan *stopword removal* adapun contoh *stopword* yang terdapat pada *stoplist library sastrawi*.

Pada tahapan ini , ada beberapa aturan proses agar hasil filtering sesuai apa yang diharapkan . aturan tersebut adalah sebagai berikut :



Tabel 4.9 Peraturan tahapan filtering

Kondisi	Aksi
Inputan data judul tugas akhir mengandung kata pada database	Maka akan menghapus kata atau string dalam data judul tugas akhir
Inputan data judul tugas akhir tidak mengandung kata pada database	Maka tidak akan dihapus kata atau string data judul tugas akhir.

Adapun isi dari sebagian daftar kata stopwords pada database yang disajikan sebagai

Tabel 4.10 Daftar sebagian kata-kata stopwords

DATA STOPWORD			
telah	punya	mendapatkan	dari
Untuk	Setiap	Pernah	Lain
Baik	Memang	Lakukan	melakukannya
Bisa	Tetapi	Sudah	Karena
Jadi	seperti	Ada	Antara
Juga	hampir	Semua	Setelah



Tabel 4.11 Di bawah ini merupakan contoh tahapan filtering sebagai berikut.

TOKENIZING	TOKENIZING	TOKENIZING	TOKENIZING	TOKENIZING
[klasifikasi] [sentimen] [terhadap] [bukalapak] [metode] [bayes] [classifier]	Aplikasi][real] [count] [studi] [kasus] [pilgub] [riau] [2018][compone nt][analysis] [pca]	[optimasi] [fuzzy] [time] [series] [menggunakan][particle] [swarm] [optimization] [untuk] [memprediksi] [kualitas] [udara] [di] [kota] [pekanbaru]	[klasifikasi] [emosi][pada] [tweet] [menggunakan] [metode support] [vector machine]	[implementasi][peraturan] [menteri] [pendayagunaan][aparatur] [negara] [dan] [reformasi] [birokrasi] [nomor] [14] [tahun][2017][pada] [system] [penilaian] [tingkat][kepuasan] [pasien] [rumah] [sakit][umum] [daerah] [arifin] [achmad]
FILTERING				
D3	D4	D5	D6	D7

<p>Hak Cipta Dilindungi Undang-undang 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber: a. Pergi untuk keperluan pribadi, keluarga, dan terbatas lainnya. b. Diperjualbelikan. 2. Dilarang menyalin, menduplikasi, menyebarkan, atau melakukan tindakan lain yang merugikan tanpa izin dari UIN Suska Riau.</p>	<p>© Hak cipta milik UIN Suska Riau</p> <p>[aplikasi] [realitas] [count] [studi] [kasus] [pilgub] [riau]</p>	<p>[optimasi] [fuzzy] [time] [series] [algoritma] [particle] [swarm] [optimization] [prediksi] [indeks] [harga] [konsumen] [ihk] [kota] [pekanbaru]</p>	<p>[klasifikasi] [emosi] [tweet] [metode] [support] [vector] [machine]</p>	<p>[implementasi] [atur] [menteri] [pendayagunaan] [aparatur] [negara] [reformasi] [birokrasi] [nomor][tahun] [system][nilai] [tingkat] [puas] [pasien][rumah] [sakit] [umum] [[daerah][arifin] [achmad]</p>
---	--	---	--	---

Keterangan :

D3 merupakan hasil *Filtering* pada judul Tugas Akhir data ke-1

D4 merupakan hasil *Filtering* pada judul Tugas Akhir data ke-2

D5 merupakan hasil *Filtering* pada judul Tugas Akhir data ke-3

D6 merupakan hasil *Filtering* pada judul Tugas Akhir data ke-4

D7 merupakan hasil *Filtering* pada judul Tugas Akhir data ke-5

4.3.4 Steaming

Langkah selanjutnya adalah *steaming* pada tahapan ini tahap mencari kata dasar dari setiap kata hasil filtering dan pengembalian bentuk kata dalam suatu representasi yang sama. Steaming merupakan salah satu tahapan dalam proses text pre-processing. Steaming bertujuan untuk mentransformasikan kata menjadi kata dasarnya (root) dengan menghilangkan semua imbuhan kata (affixes) meliputi awalan kata (prefixes), sisipan kata



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip, sebagian atau seluruhnya atau cara apa pun untuk disebarluaskan dalam bentuk cetak, elektronik, mekanis, fotokopi, rekam, atau dengan cara lain tanpa mengizinkan penulis atau penerbit untuk memperbanyak atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
2. Dilarang menguraikan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

akhiran kata (suffixes) dan atau menghilangkan awalan dan akhiran kata (prefixes) pada kata turunan.

Salah satu library yang bisa digunakan dalam melakukan proses stemming bahasa Indonesia adalah dengan menggunakan Library Python Sastrawi . Library ini merupakan turunan dari Library PHP Sastrawi dimana library tersebut menerapkan algoritma Nazief dan Adriani . Tahapan algoritma tersebut meliputi :

Langkah pertama adalah memeriksa apakah kata tersebut merupakan akar kata (root) terdapat dalam daftar akar kata (root). Kita kata tersebut merupakan akar kata, maka proses dihentikan pada tahap pertama ini.

Menghilangkan Inflection Suffixes (-lah , -kah , -ku , -mu , atau -nya). Jika kata berupa particles (-lah , -kah , -tah atau -pun) maka langkah ini diulangi lagi untuk menghapus Possesive Pronouns (-ku , -mu , atau -nya).

Menghilangkan derivational suffix (imbuhan turunan). Hilangkan imbuhan -i, -kan, an.

Menghilangkan derivational prefix (awalan turunan). Hilangkan awalan be-, di-, ke-me-, pe-, se- dan te-.

Bila dari langkah 4 di atas belum ketemu juga. Maka lakukan analisis apakah kata tersebut masuk dalam tabel diambiguitas kolom terakhir atau tidak.

Bila dari langkah 4 di atas belum ketemu juga. Maka lakukan analisis apakah kata tersebut masuk dalam tabel diambiguitas kolom terakhir atau tidak.

Bila semua proses di atas gagal, maka algoritma mengembalikan kata aslinya.



Tabel 4.12 kombinasi awalan akhiran yang dilarang

Awalan (prefix)	Akhiran (suffix) yang dilarang
	-i
	-an
	-i,-kan
	-an
	-i,-kan
	-an

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya tulis, naskah, atau seluruhnya. b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau. c. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 4.13 Aturan pemenggalan awalan

Aturan	Format Kata	Pemenggalan
1	berV...	ber-V... be-rV...
2	berCAP...	ber-CAP... dimana C!="r" & P!="er"
3	berCAerV...	ber-CaerV... dimana C!="r"
4	Belajar	bel-ajar
Aturan	Format Kata	Pemenggalan
5	beC1erC2...	be-C1erC2... dimana C1!={"r"} {"1"}
6	terV...	ter-V... te-rV...
7	terCerV...	ter-CerV... dimana C!="r"
8	terCP...	ter-CP... dimana C!="r" dan P!="er"
9	teC1erC2...	te-C1erC2... dimana C1!="r"
10	me {l r w y} V...	me- {l r w y} V...
11	mem {b f v} ...	mem- {b f v} ...
12	mempe...	mem-pe...
13	mem {rV V} ...	me-m {rV V} ... me-p {rV V} ...
14	men {c d j z} ...	men- {c d j z} ...

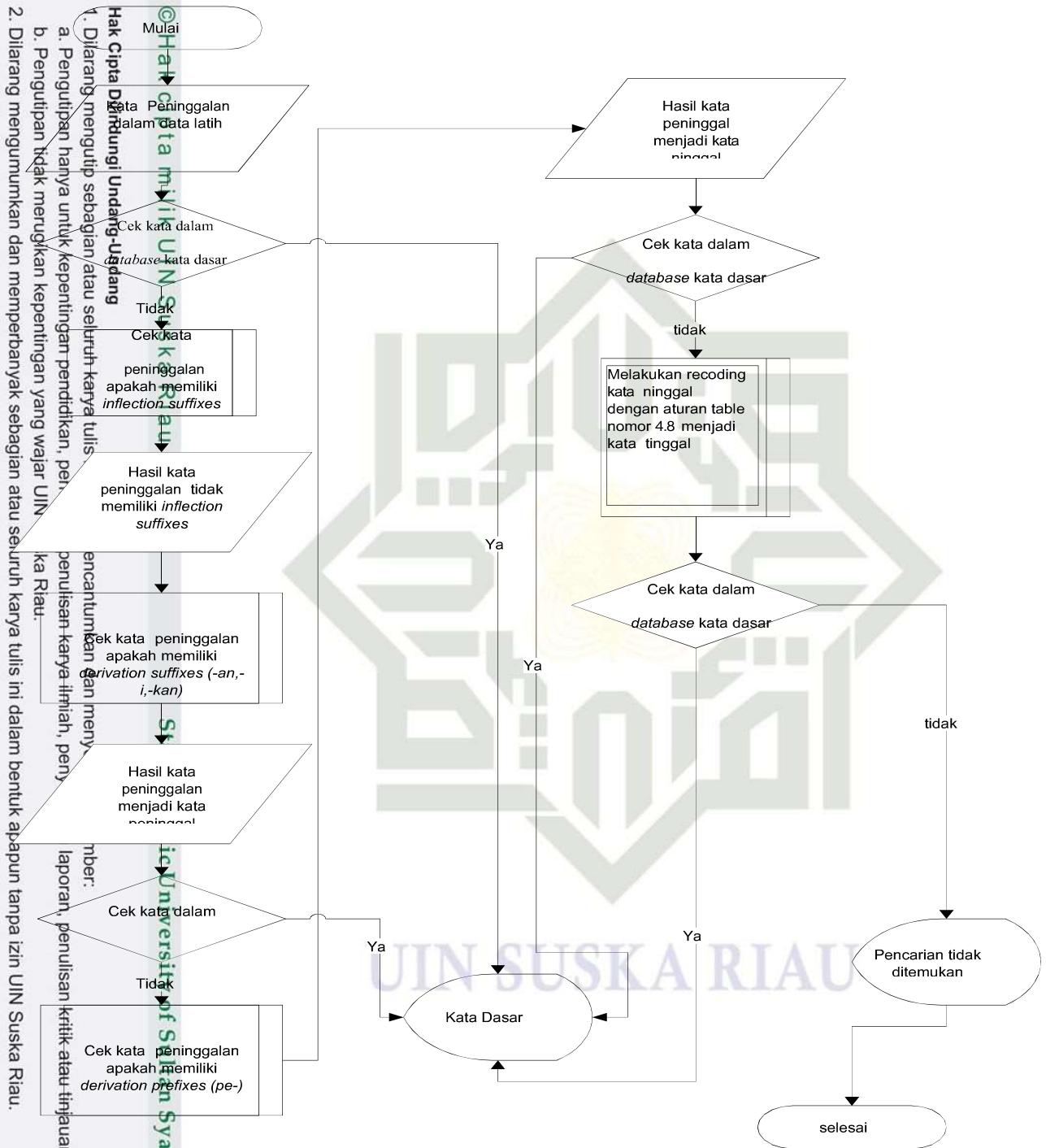


2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

15	menV...	me-nV... me-tV
16	meng {g h q k}...	meng- {g h q k}...
18	mengV...	meng-V... meng-kV...
18	menyV...	meny-sV...
19	mempV...	mem-pV... dengan V!="e"
20	pe {w y} V...	pe- {w y} V...
21	perV...	per-V... pe-rV...
22	perCAP	per-CAP... dimana C!="r" dan P!="er"
22	perCAerV...	per-CAerV... dimana C!="r"
23	pem {b f V} ...	pem- {b f V} ...
23	pem {rV V} ...	pe-m {rV V} ... pe-p {rV V} ...
24	pen {c d j z} ...	pen- {c d j z} ...
24	penV...	pe-nV... pe-tV...
25	peng {g h q}...	peng- {g h q}...
25	pengV...	peng-V... peng-kV...
30	penyV...	peny-sV...
31	peIV...	pe-IV... kecuali pelajar yang menghasilkan ajar
31	peCerV...	per-erV... dimana C!={r w y l m n}
31	peCP...	pe-CP... dimana C!={r w y l m n} dan P!="er"
33	terC1erC2...	ter-C1erC2... dimana C1!="r"
35	peC1erC2...	pe-C1erC2... dimana C1!={r w y l m n}

Pada hasil *filtering* masih ada beberapa kata yang memiliki imbuhan, maka kata-kata tersebut harus diproses dengan menggunakan algoritma *stemming nazief dan Adriani*. Proses *stemming* ini mempunyai alur yang digambarkan pada gambar 4.6 sebagai berikut.:



Gambar 4.6 flowchart alur proses *steaming*



1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan sumber.
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Pada contoh kasus hanya mengambil kata Peninggalan . Tahapan yang dilakukan untuk kata Peninggalan , dikerjakan sebagai berikut :

Kata yang hendak di *stemming* dicari terlebih dahulu pada kamus. Jika ditemukan, berarti kata tersebut adalah kata dasar, jika tidak maka langkah 2 yang dilakukan.

Pada kata Peninggalan , akan dicek apakah memiliki *inflection suffixes*, yaitu akhiran(-lah , -kah , -tah , -pun) dan kata ganti kepemilikan atau *possessive pronoun* PP (-ku , -mu , -nya) ternyata pada kata Peninggalan tidak terdapat *inflectional particle* maka proses selanjutnya.

Hilangkan *Derivation Suffixes* DS (-i , -kan , atau -an). Ternyata kata peninggalan terdapat *Derivation Suffixes* dimana terdapat akhiran -an maka hapus akhiran -an , sehingga kata menjadi peninggal , kemudian sistem mengerjakan pencarian ke *database* jika kata peninggal ditemukan maka proses berhenti jika tidak ditemukan akan dilakukan proses selanjutnya.

Kata peninggal akan dicek, apakah memiliki *derivation prefixes*, ternyata kata peninggal mengandung *derivation prefixes* -pe, sehingga kata menjadi ninggal , kemudian sistem mencari kata ninggal ke database jika ditemukan maka kata ninggal adalah kata dasar dan proses berhenti, jika tidak maka kata ninggal kembali menjadi kata peninggal dan melanjutkan proses berikutnya.

Setelah selesai menjalankan langkah 1 sampai 4, tetapi masih belum menemukan kata dasar maka tahapan selanjutnya adalah *recoding* menggunakan aturan pada tabel 4.8 Kata peninggal merupakan kata yang memiliki imbuhan peng- dan diikuti huruf vokal, dengan kondisi tersebut maka aturan yang dipakai adalah aturan nomor 27. Pertama sistem akan memotong kata dengan pen- maka hasilnya menjadi kata inggal selanjutnya sistem akan mencari kata inggal ke database jika ditemukan maka kata inggal adalah kata dasar jika tidak maka sistem akan memenggal kata peninggal dengan pen- dan



1. Dilarang menyalin, mengutip, atau menjiplak sebagian atau seluruh isi dari dokumen ini tanpa izin tertulis dari penerbit.
 2. Dilarang menguraikan, menyalin, mengutip, atau menjiplak sebagian atau seluruh isi dari dokumen ini dalam bentuk apapun tanpa izin UIN Suska Riau.

menambah huruf t pada hasil pemenggalan sehingga kata peninggal menjadi tinggal, selanjutnya sistem mencari kata tinggal ke database dan ternyata kata tinggal terdapat dalam *database*, maka proses berhenti.

Sehingga merupakan contoh tahapan *stemming* pada data judul Tugas Akhir dapat dilihat sebagai berikut.

Tabel 4.13 Tabel *Steaming*

D3	D4	D5	D6	D7
[aplikasi]	[real]	[optimasi]	[klasifikasi]	[implementasi]
[klasifikasi]	[count]	[fuzzy] [time]	[emosi]	[atur] [menteri]
[sentimen]	[studi]	[series]	[tweet]	[pendayagunaan]
[hadap]	[kasus]	[algoritma]	[metode]	[aparatur]
[bukalapak]	[pilgub]	[particle]	[support]	[negara]
[metode]	[riau]	[swarm]	[vector]	[reformasi]
[na] [ve]		[optimization]	[machine]	[birokrasi]
[bayes]		[prediksi]		[nomor] [tahun]
[classifier]		[indeks]		[sistem] [nilai]
		[harga]		[tingkat] [puas]
		[konsumen]		[pasien] [rumah]
		[ihk] [kota]		[sakit] [umum]
		[pekanbaru]		[daerah] [arifin]
				[achmad]



© Hikmah Iqbal, UIN Suska Riau
 Dilarang mengumpukan dan memperbanyak sebagian atau seluruhnya tanpa izin UIN Suska Riau.

© Hikmah Iqbal, UIN Suska Riau
 Dilarang mengumpukan dan memperbanyak sebagian atau seluruhnya tanpa izin UIN Suska Riau.

- Keterangan :
- 1. D merupakan hasil *steaming* kata (Mengggunakan) menjadi (Guna) pada judul Tugas Akhir data ke-1
 - 2. D merupakan hasil *steaming* kata (Mengggunakan) menjadi (Guna) pada judul Tugas Akhir data ke-2
 - 3. D merupakan hasil *steaming* pada judul Tugas Akhir data ke-3
 - 4. D merupakan hasil *steaming* kata (Mengggunakan) menjadi (Guna) pada judul Tugas Akhir data ke-4
 - 5. D7 merupakan hasil *steaming* pada judul Tugas Akhir data ke-5

4.3. TF-IDF (Term Frequency – Inverse Document Frequency)

Algoritma TF-IDF (Term Frequency – Inverse Document Frequency) adalah salah satu algoritma yang dapat digunakan untuk menganalisa hubungan antara sebuah frase/kalimat dengan sekumpulan dokumen.

4.4. Term Frequency

Term frequency bertujuan untuk menghitung kemunculan suatu *term* berdasarkan bobot suatu *term* pada dokumen tertentu. Dalam suatu dokumen apabila *term* tertentu memiliki kemunculan yang tinggi, maka akan semakin tinggi bobot dokumen untuk *term* tersebut dan sebaliknya.

Pada *Term frequency* (TF), terdapat beberapa jenis formula yang dapat digunakan diantaranya :

- 1. TF Biner (*Binary TF*)
 Hanya memperhatikan apakah suatu kata atau *term* ada atau tidak dalam dokumen. jika ada diberi nilai satu (1) jika tidak diberi nilai nol (0).
- 2. TF Murni (*raw TF*)



Nilai Tf diberikan berdasarkan jumlah kemunculan suatu *term* di dalam dokumen. Contohnya jika muncul lima (5) kali maka kata tersebut akan bernilai lima (5).
 Logaritmik ini untuk menghindari dominasi dokumen yang mengandung sedikit *term* dalam *query*, namun mempunyai frekuensi yang tinggi.

Tabel Tf data Judul Tugas akhir

Term	Tf					
	D4	D3	D5	D6	D7	df
emosi	1					1
tweet	1					1
support	1					1
vector	1					1
mechine	1					1
klasifikasi		1		1		2
sentimen		1				1
hadap		1				1
bukalapak		1				1
metode		1		1		2
naive		1				1
bayes		1				1
clasifier			1			1
aplikasi			1			1
real			1			1
count			1			1
studi			1			1
kasus			1			1

1. Tidak boleh menjiplak atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

			1			1
pilgub			1			1
riau			1			1
optimasi			1			1
fuzzi			1			1
time			1			1
series			1			1
algorima			1			1
particle			1			1
swarm			1			1
optimization				1		1
prediksi				1		1
indeks				1		1
harga				1		1
konsumen				1		1
ink				1		1
kota				1		1
pekanbaru				1		1
implementasi					1	1
atur					1	1
negara					1	1
reformasi					1	1
birokrasi					1	1
nomor					1	1
tahun					1	1
sistem					1	1
nilai					1	1



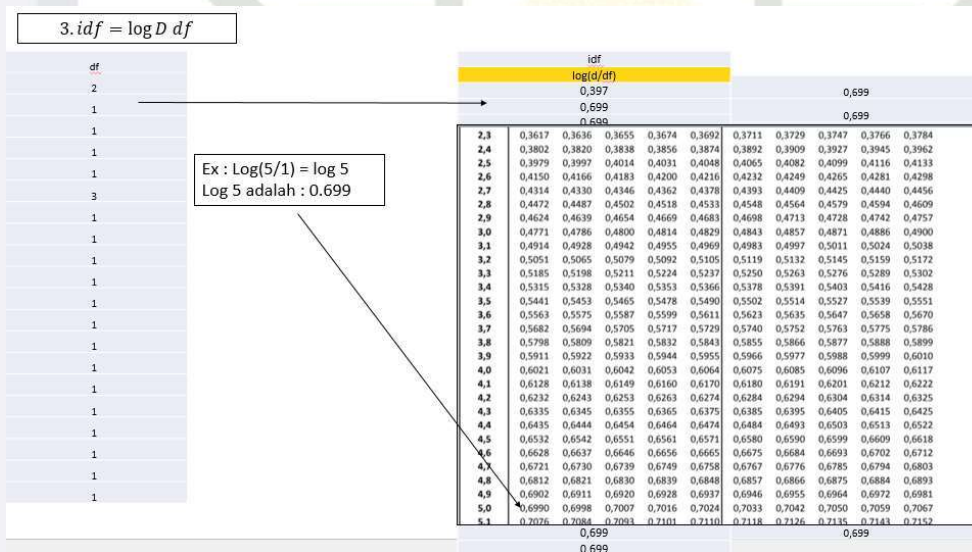
Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis atau tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpukan dan memperbanyak atau memperjualbelikan karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

tingkat				1	1
puas				1	1
pasien				1	1
rumah				1	1
sakit				1	1
umum				1	1

4.2 Inverse Document Frequency (IDF)

Inverse document frequency (IDF) merupakan perhitungan dari bagaimana setiap kata ini di susun secara luas pada koleksi dokumen yang bersangkutan. IDF juga menunjukkan jumlah keseterediaan sebuah kata dengan seluruh dokumen. Semakin sedikit jumlah dokumen yang mengandung kata yang dimaksud maka nilai IDF semakin besar. Berikut adalah perhitungan IDF:



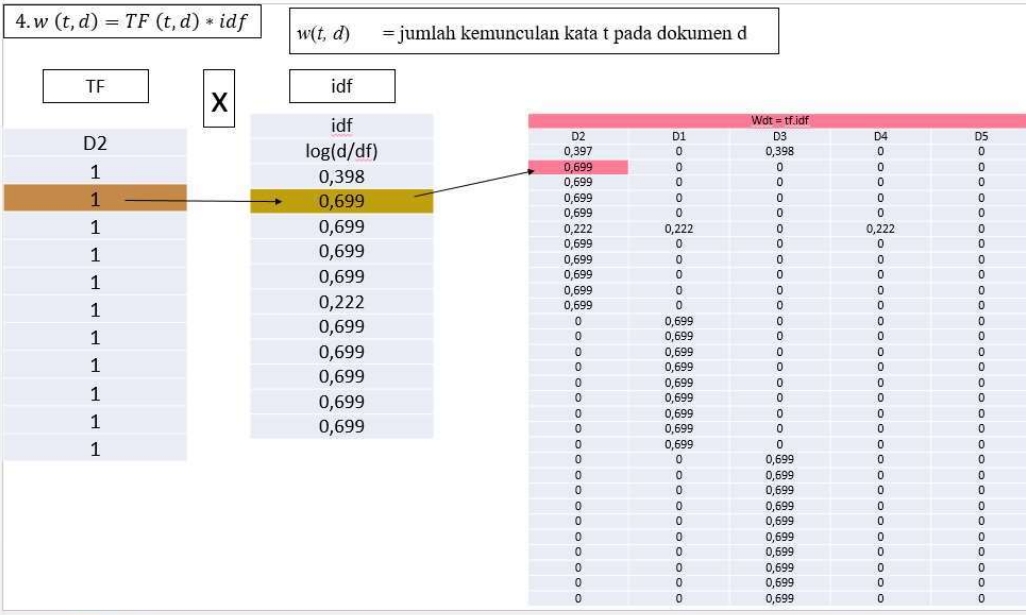
Gambar 4.7 perhitungan IDF

pada contoh diatas $\log 5 = 0,699$ dapat dilihat dari tabel logaritma, $1 = df$, jadi untuk hasil $idf = 0,699/1 = 0,699$



Setelah didapatkan hasil Tf dan Df maka dapat dilihat pada contoh gambar berikut untuk

proses perkalian hasil dari TF*IDF, dapat dilihat sebagai berikut :



Gambar 4.8 Hasil perkalian TF-IDF

Pada perhitungan TF-idf yang merujuk pada rumus 2.1 yang dapat diperoleh dari contoh data untuk tugas akhir maka dapat dilihat dari tabel berikut ini :

Tabel 4.16 Perhitungan TF-IDF

idf	Wdt = tf.idf				
	D4	D3	D5	D6	D7
log(n/df)					
0,699	0,699	0	0,000	0	0
0,699	0,699	0	0	0	0
0,699	0,699	0	0	0	0
0,699	0,699	0	0	0	0
0,699	0,699	0	0	0	0



© Hak cipta milik UIN Suska Riau State Islamic University of Sultan Syarif Kasim Riau

0,000	0,398	0	0,398	0
0,000	0,698970004	0	0	0
0,000	0,698970004	0	0	0
0,000	0,698970004	0	0	0
0,000	0,397940009	0	0,397940009	0
0,000	0,698970004	0	0	0
0	0,699	0	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0,698970004	0	0
0	0,000	0	0,698970004	0
0	0,000	0	0,698970004	0
0	0,000	0,000	0,698970004	0

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
- 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



0	0,000	0	0,698970004	0
0	0,000	0	0,698970004	0
0	0,000	0	0,698970004	0
0	0,000	0	0,698970004	0
0	0,000	0	0,698970004	0
0	0,000	0	0	0,698970004
0	0,000	0	0	0,698970004
0	0	0	0	0,698970004
0	0	0	0	0,698970004
0	0	0	0	0,698970004
0	0	0	0	0,698970004
0	0	0	0	0,698970004
0	0	0	0	0,698970004
0	0	0	0	0,699
0	0	0	0	0,699
0	0	0	0	0,699
0	0	0	0	0,699
0	0	0	0	0,699
0	0	0	0	0,699
0	0	0	0	0,699
0	0	0	0	0,699

1. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Keterangan

1. Diberikan hasil perkalian banyak TF pada dokumen D2 dikali idf D3.

2. Diberikan hasil perkalian banyak TF pada dokumen D1 dikali idf D4.

3. Diberikan hasil perkalian banyak TF pada dokumen D3 dikali idf D5.

4. Diberikan hasil perkalian banyak TF pada dokumen D4 dikali idf D6.

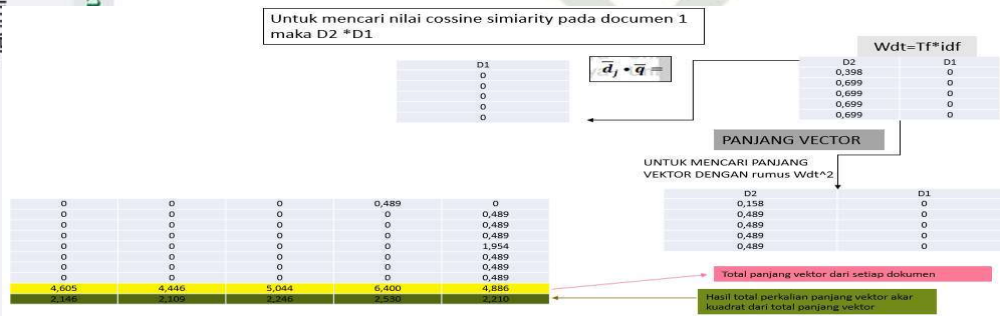
5. Diberikan hasil perkalian banyak TF pada dokumen D5 dikali idf D7.

2.1. Klasifikasi Algoritma NWKNN

2.1.1. Cosine Similarity

Cosine similarity merupakan ukuran keamaan salah satu implementasi untuk mencari tingkat kemiripan teks pada teks itu sendiri atau pada sebuah kalimat pada dua buah dokumen atau lebih.

Cosine similarity berfungsi untuk membandingkan kemiripan antar dokumen, dalam hal ini yang dibandingkan adalah query dengan dokumen latih., setelah itu di hitung akar pangkat dua. Gambar berikut merupakan contoh dari perhitungan *cosine similarity* merujuk pada rumus 2.3 :



Gambar 4.9 contoh perhitungan *cosine similarity*



Pada data tugas akhir Dalam menghitung cosine similarity:

pertama yang di lakukan yaitu melakukan perkalian skalar antara query dengan dokumen kemudian dijumlahkan. Pada tabel diatas dapat dijelaskan bahwa yang menjadi queri disini D4 karena yang diklasifikan adalah D4 (Data ke-304) dan yang sebagai dokumen adalah D3,D5,D6,D7. Misal $D4 * D3 = D3$, $D4 * D5 = D4$, $D4 * D6 = D5$, $D4 * D6 = D6$.

Tabel 4.17 Hasil Wdt(jumlah kemunculan t pada data d)

Idf	Wdt = tf.idf				
	D4	D3	D5	D6	D7
0,699	0,699	0	0,000	0	0
0,699	0,699	0	0	0	0
0,699	0,699	0	0	0	0
0,699	0,699	0	0	0	0
0,699	0,699	0	0	0	0
0,699	0,699	0	0	0	0
0,000	0,000	0,398	0	0,398	0
0,699	0,000	0,698970004	0	0	0
0,699	0,000	0,698970004	0	0	0
0,699	0,000	0,698970004	0	0	0
0,000	0,000	0,397940009	0	0,397940009	0
0,699	0,000	0,698970004	0	0	0
0,699	0	0,699	0	0	0
0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0,698970004	0	0

Hak Cipta dilindungi undang-undang. Dilarang mengutip, menyebarkan, atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

State Islamic University of Sultan Syarif Kasim Riau



2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0,698970004	0	0
0,699	0	0,000	0	0,698970004	0
0,699	0	0,000	0	0,698970004	0
0,699	0	0,000	0,000	0,698970004	0
0,699	0	0,000	0	0,698970004	0
0,699	0	0,000	0	0,698970004	0
0,699	0	0,000	0	0,698970004	0
0,699	0	0,000	0	0,698970004	0
0,699	0	0,000	0	0,698970004	0
0,699	0	0,000	0	0,698970004	0
0,699	0	0,000	0	0,698970004	0
0,699	0	0,000	0	0	0,698970004
0,699	0	0,000	0	0	0,698970004
0,699	0	0	0	0	0,698970004
0,699	0	0	0	0	0,698970004
0,699	0	0	0	0	0,698970004
0,699	0	0	0	0	0,698970004
0,699	0	0	0	0	0,699
0,699	0	0	0	0	0,699
0,699	0	0	0	0	0,699
0,699	0	0	0	0	0,699



0	0	0	0	0,489
0	0	0	0	0,489
0	0	0	0	0,489
0	0	0	0	0,489
2,443	2,760	7,328	4,225	7,328

setelah itu di hitung akar pangkat dua pada tabel berikut adalah hasil jumlah dari akar pangkat dua dari panjang vektor

2,146	2,109	2,246	2,530	2,210
-------	-------	-------	-------	-------

Selanjutnya hasil perkalian skalar tersebut di bagi dengan hasil perkalian panjang dokumen dan query. Tabel berikut menunjukkan hasil akhir dari *cosine similarity*.

Tabel 4.20 Hasil cosine similarity

D1	D3	D4	D5
0,011	0,000	0,033	0,000
CS	CS	CS	TI

4.5.2 Mengurutkan Hasil Kedekatan K Tetangga

Pengurutan dilakukan mulai dari nilai tertinggi sampai nilai terendah, seperti terlihat pada tabel 4.21 dibawah ini.

Tabel 4.21 Pengurutan Hasil Kedekatan K Tetangga

Hasil Cossim	Kelas
0,033	CS
0,011	CS
0,009	CS



Hak Cipta Dilindungi Undang-Undang
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber.
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penerjemahan atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengutamakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

0,000	TI
-------	----

4.2 Pembobotan Setiap Kelas

Pada pembobotan ini dicontohkan K=3 dan E=2. K=3 adalah kedekatan ketetanggan ke-1 sampai data ke-3, sedangkan untuk E=2 merupakan nilai exponen dengan ketentuan 1/Exponen Hasilnya Dapat dilihat pada tabel 4.20 dibawah ini.

Tabel 4.22 hasil bobot tiap kelas

Hasil Cossim	Kelas
0,033	CS
0,011	CS
0,009	CS

Tabel 4.23 Tabel kelas tertinggi dan minimum

Kelas	Jumlah
CS	3
TI	1

Kemudian hitung bobot pada tiap kelas dengan 1 dikali mengambil nilai kelas tertinggi dibagi dengan nilai kelas minimum kemudian dibagi 1/jumlah E(Exponen) Sehingga didapat nilai bobot tiap kelas sebai berikut :

Tabel 4.24 Hasil nilai bobot tiap kelas

Kelas	Bobot
CS	0,577350269
TI	0



4.5.4 Perhitungan skor

telah mendapat nilai bobot, maka menghitung nilai skor dengan cara mengalikan kelas dan jumlah skor dari tiap kelas dengan masing-masing bobot dari tiap kelas. Untuk menghitung skor dapat menggunakan dengan ketentuan nilai : Bernilai 1 jika dengan sesama jenis bernilai 0 jika tidak sesama jenis. Dapat dilihat pada tabel berikut :

Tabel 4.25 Ketentuan skor

Klasifikasi		Nilai
CS	CS	1
CS	TI	0
TI	CS	0
TI	TI	1

Tabel 4.26 Perhitungan Skor

mengambil data yang telah diurutkan pada K = 3

Hasil Cossim	Kelas
0,033	CS
0,011	CS
0,009	CS

Dari tabel diatas dijelaskan bahwa bobot pada tabel 4.23 dikali hasil cossine pertama dikali dengan ketentuan skor ditambah hasil cossine kedua dikali dengan ketentuan skor, ditambah hasil cossine ketiga dikali dengan ketentuan nilai skor. Maka dapat dilihat pada tabel berikut hasil bobot dengan perhitungan skor.

Skor(CS) =	0,030
------------	-------



Masing-masing nilai bobot yang didapatkan untuk masing-masing kelas digunakan untuk menghitung nilai skor pada data uji terhadap setiap kelas yang terdefinisi pada data latih. Skor kelas dengan nilai tertinggi akan menjadi kelas klasifikasi data uji.

Hasil klasifikasi data judul tugas akhir pada 412 record data dapat ditunjukkan pada gambar berikut :

No	Judul	TI	CS	Hasil Klasifikasi	Label	Hasil
1	KLASIFIKASI EMOSI PADA TWEET MENGGUNAKAN METODE SUPPORT VECTOR MACHINE	0	6769.784	CS	CS	True
2	KLASIFIKASI TWEET PADA TWITTER BERDASARKAN MINAT DENGAN METODE SUPPORT VEKTOR MACHINE (SVM)	0	32.614	CS	CS	True
3	KLASIFIKASI KOMENTAR PADA PLAY STORE DENGAN MENGGUNAKAN METODE MODIFIED K-NEAREST NEIGHBOR (MKNN)	0	1409.351	CS	CS	True
4	ANALISA SENTIMEN TWITTER TERHADAP LAYANAN OVO MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER	0	1228.684	CS	CS	True
5	ANALISIS SENTIMEN BODY SHAMING PADA KOMENTAR INSTAGRAM MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER	0	1362.166	CS	CS	True
6	KLASIFIKASI SENTIMEN TERHADAP BUKALAPAK DENGAN MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER	0	1472.742	CS	CS	True

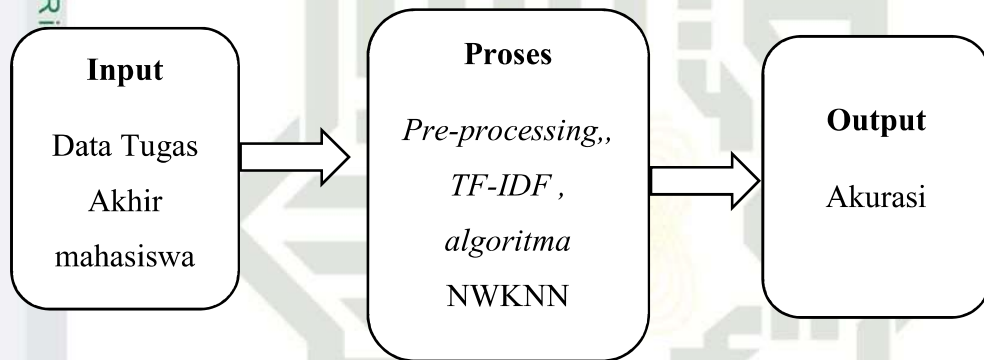
Gambar 4.10 Hasil Klasifikasi Metode NWKNN

Pada gambar diatas diperoleh hasil tertinggi untuk perhitungan skor data judul tugas akhir pada baris judul data tugas akhir data ke-304 dengan skor 6769.784 dengan hasil klasifikasi *computer science*.



4.7 Perancangan Umum

Perancangan umum merupakan susunan atau rancangan algoritma yang dibangun. Perancangan ini bertujuan untuk mewujudkan perangkat lunak sesuai dengan analisa kebutuhan yang telah ditetapkan sebelumnya, proses perancangan umum untuk menemukan akurasi dari data Tugas Akhir yang memiliki 2 bidang keahlian dijelaskan pada gambar berikut.



Gambar 4.11 Diagram menemukan akurasi

a. Perancangan Interface

Perancangan antar muka merupakan perancangan aplikasi yang digunakan sebagai acuan dalam implementasi aplikasi yang akan dibangun. Perancangan antarmuka bertujuan untuk memudahkan pengguna dalam menggunakan aplikasi.

b. Rancangan Dashboard

Tampilan ini merupakan tampilan yang menampilkan halaman awal aplikasi.



1. Hi

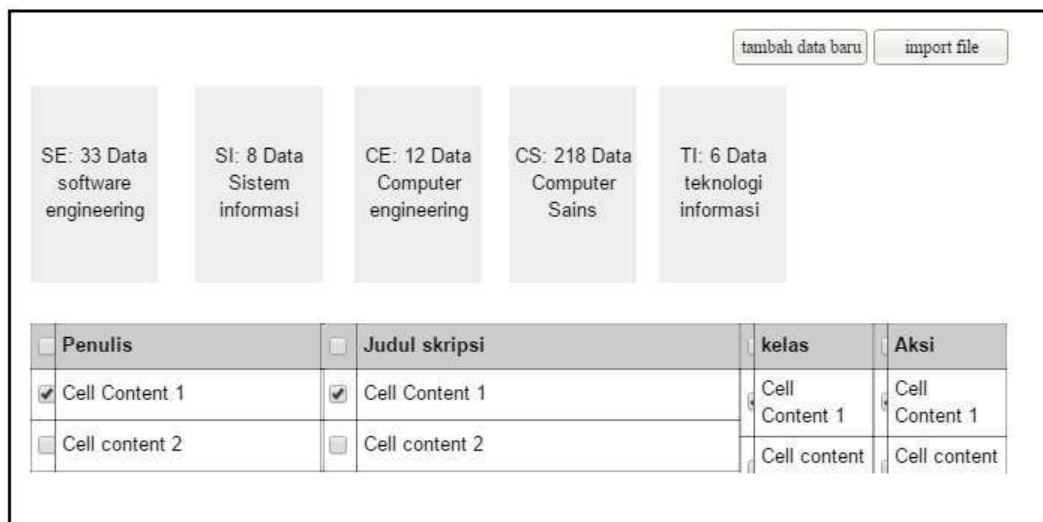
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruhnya tanpa izin UIN Suska Riau.



Gambar 4.12 Rancangan Halaman Awal

2. Rancangan Menu Data Tugas Akhir

Tahapan ini merupakan rancangan pada *menu proses input* data Tugas Akhir yang digunakan sebagai daftar data Tugas Akhir yang akan dilakukan klasifikasi. Rancangan menu dapat dilihat pada gambar dibawah ini ;



Gambar 4.13 Input Data Tugas Akhir

if Kasim Riau
suatu masalah.



2. Dilarang menguraikan...
3. Rancangan Menu Pembagian Data
- a. Pengujian hanya u...
- b. Pengujian tidak me...

State Islamic University of Sultan Syarif Kasim Riau

menyebutkan sumber...

lah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

Unigram.

Jalam bentuk apapun tanpa izin UIN Suska Riau.

Pada menu ini merupakan menu yang menampilkan pembagian data latih dan data uji untuk tahap dilakukannya perhitungan akurasi pada klasifikasi data Tugas Akhir yang dapat dilihat pada gambar berikut :

	Data latih		Data Uji
<input checked="" type="checkbox"/>	70%	<input checked="" type="checkbox"/>	30%
<input type="checkbox"/>	60%	<input type="checkbox"/>	40%
<input checked="" type="checkbox"/>	50%	<input checked="" type="checkbox"/>	50%

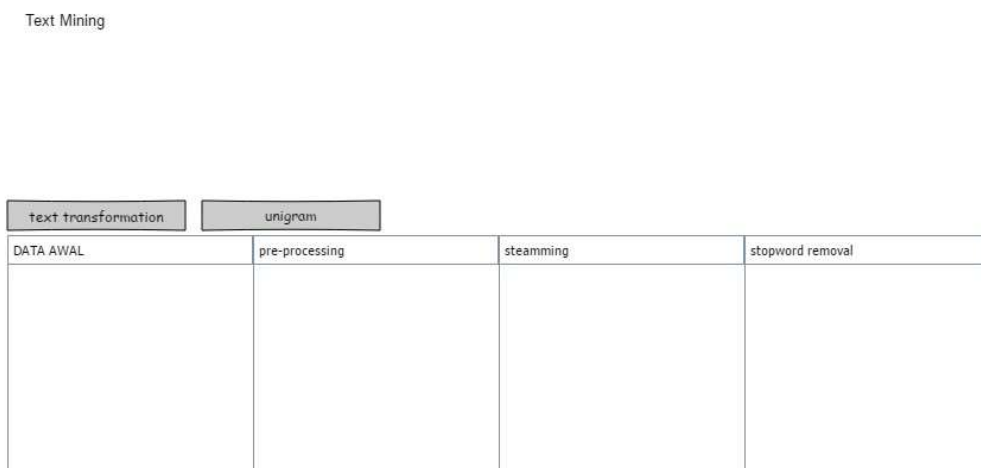
Gambar 4.14 Pembagian Data Latih Dan Uji

4. Rancangan Menu Text mining

Pada menu ini merupakan menu yang menampilkan proses text mining pada data tugas Akhir diantaranya terdapat proses pre processing ,stemming, stopword removal,

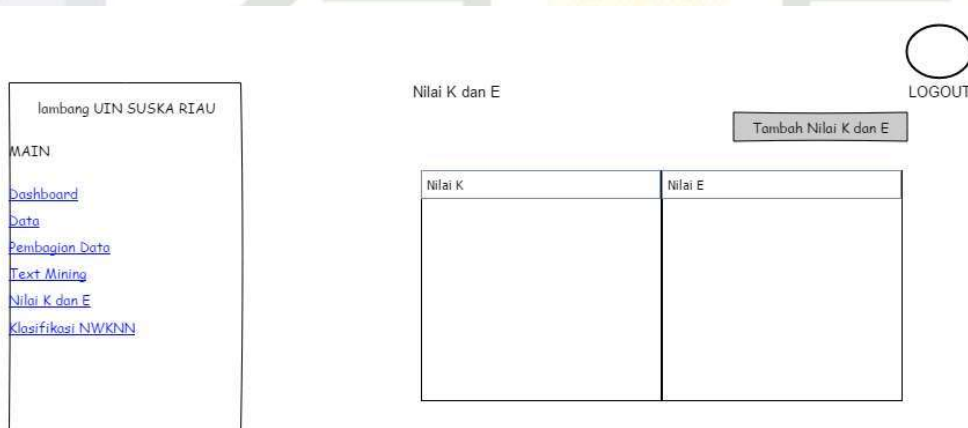
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.



Gambar 4.15 Proses *Text Processing*

Rancangan Menu nilai K dan E



Gambar 4.16 Menu Nilai K Dan E

Rancangan Menu Klasifikasi NWKNN

Pada *menu klasifikasi NWKNN* merupakan *menu* proses pengklasifikasian data tugas Akhir yang telah diberi label berdasarkan 5 bidang keahlian. Pada



lambang UIN SUSKA RIAU

MAIN

[Dashboard](#)

[Data](#)

[Pembagian Data](#)

[Text Mining](#)

[Nilai K dan E](#)

[Klasifikasi NWKNN](#)

klasifikasi NWKNN



DATA LATIH

70

NILAI K dan E

k=4 dan E=12

proses

Gambar 4.17 Proses Klasifikasi NWKNN



BAB V

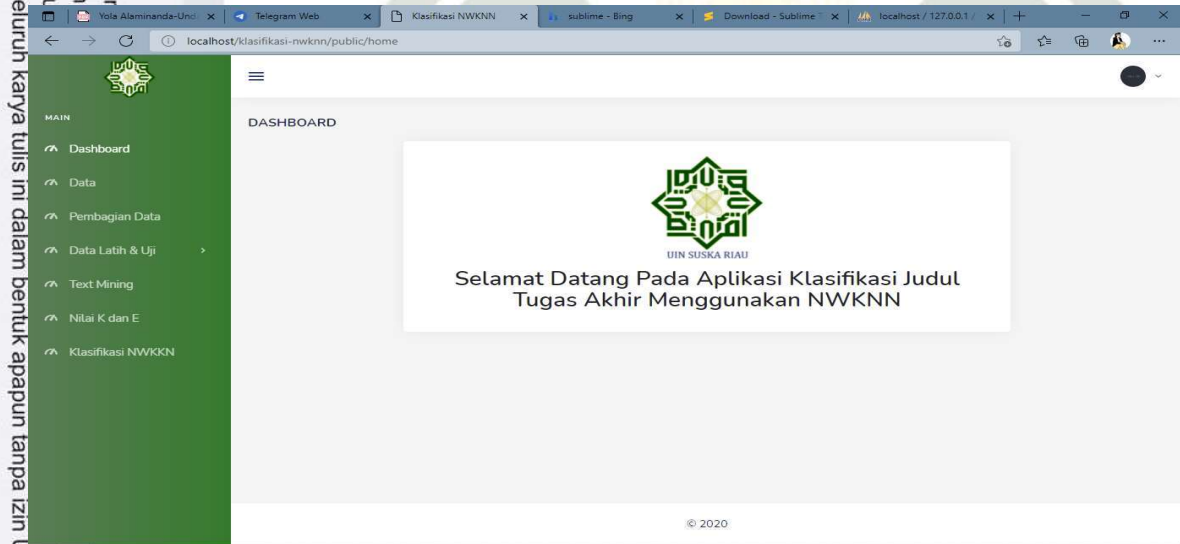
IMPLEMENTASI DAN PENGUJIAN

Tahapan implementasi dan pengujian merupakan tahap yang dilakukan setelah analisis dan perancangan. Implementasi menjelaskan mengenai penerapan dari sebuah aplikasi yang telah dianalisa dan dirancang sebelumnya. Pengujian dilakukan untuk menjamin dan menguji bahwa aplikasi yang telah dibangun sesuai dengan tujuan sebelumnya.

5. Implementasi *Interface* (Antarmuka)

1. Tampilan *Dashboard*

Tampilan ini merupakan tampilan yang menampilkan halaman awal aplikasi.

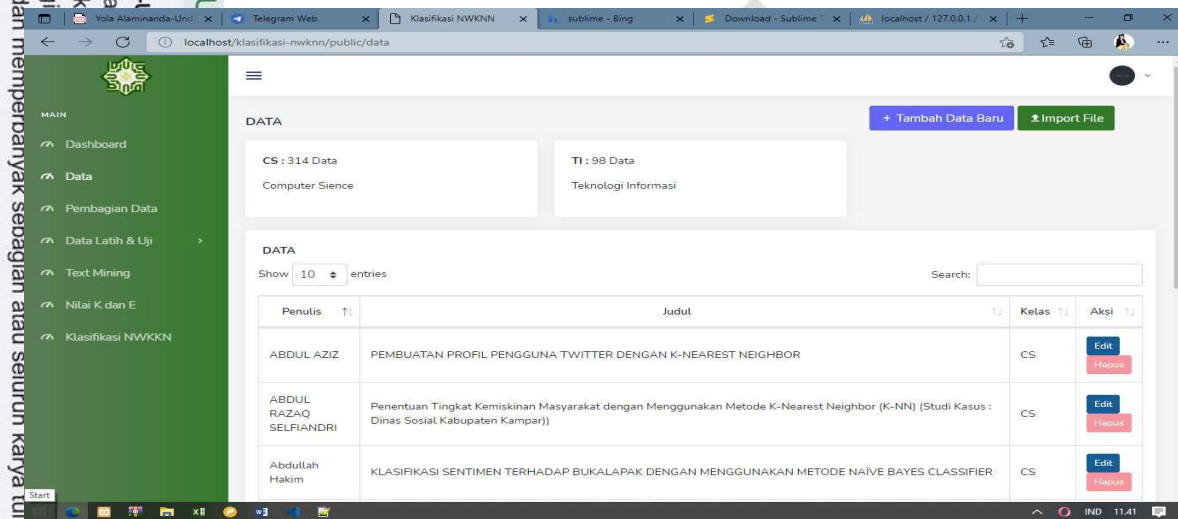


Gambar 5.1 Tampilan *Dashboard System*



2. Tampilan Menu Data

Tampilan ini merupakan tampilan yang menampilkan Data yang terdapat pada aplikasi yang akan diklasifikasikan. pada menu ini dapat ditambahkan data judul tugas Akhir baru dengan manual atau *format excel* Tampilan ini dapat dilihat dari gambar berikut ini :



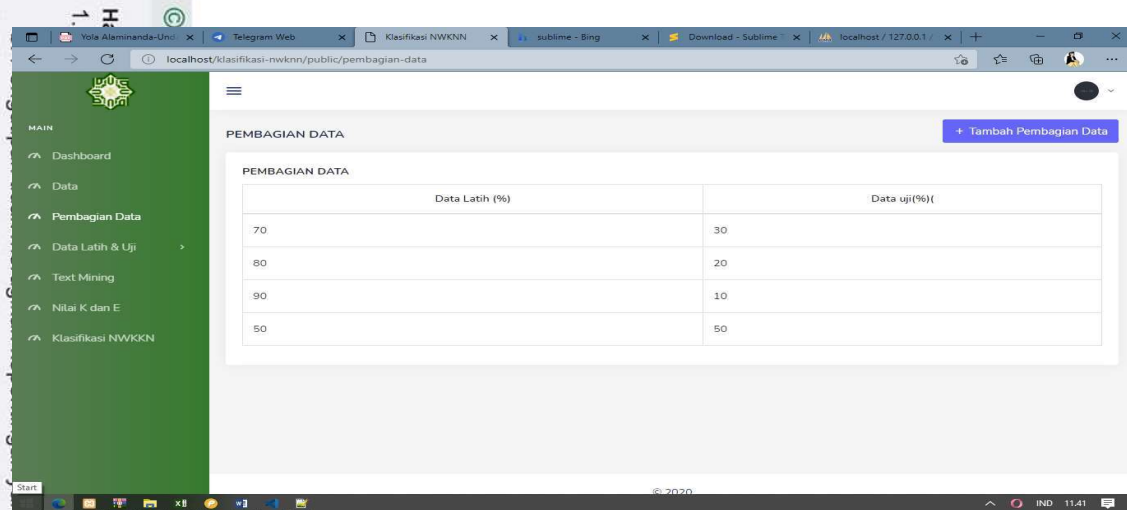
Gambar 5.2 Tampilan Menu Data System

3. Tampilan Menu Pembagian Data

Tampilan ini merupakan tampilan yang menampilkan pembagian data pada data tugas akhir terhadap data uji dan data latih yang akan dijadikan untuk tahap pengklasifikasian pada data judul tugas akhir dengan metode NWKNN. Tampilan menu pembagian data ini dapat dilihat dari gambar berikut ini :



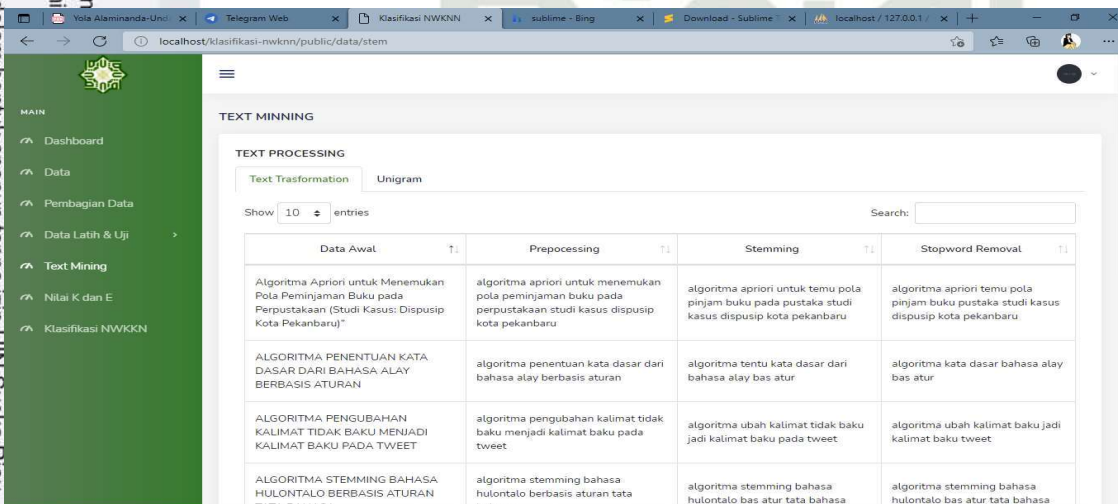
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 5.3 Tampilan Menu Pembagian Data System

4. Tampilan Menu *Text Mining*

Tampilan ini merupakan tampilan yang menampilkan proses terjadinya *text mining* pada data Tugas Akhir yang akan diklasifikasikan dengan metode NWKNN. Tampilan proses *text mining* ini dapat dilihat dari gambar berikut ini :



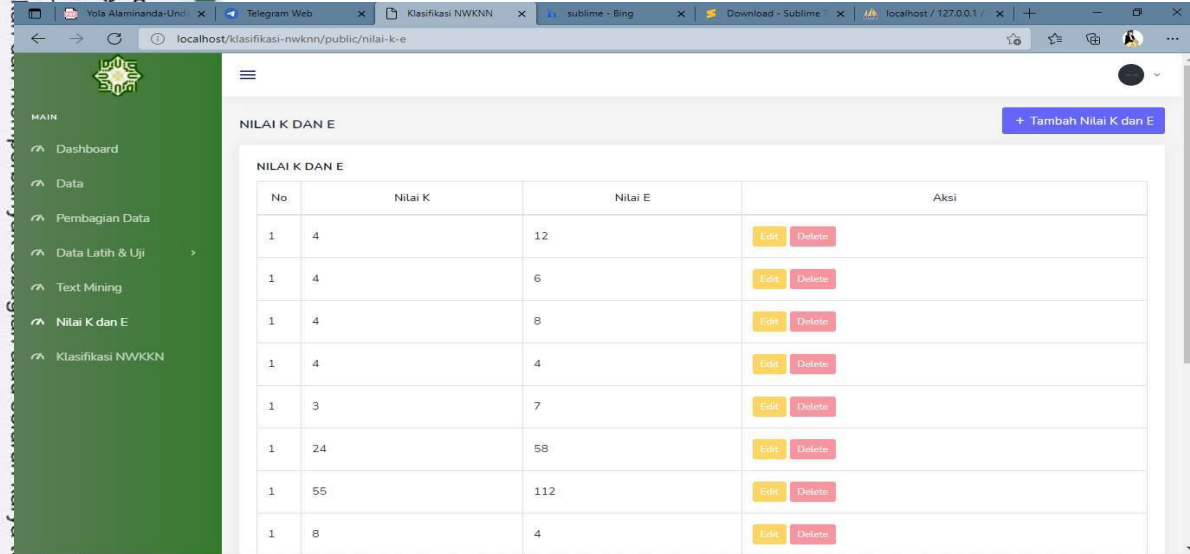
Gambar 5.4 Tampilan Menu *Text mining System*



2. Dilarang mengunm
 5. Tampilan Menu Nilai K dan E
 1. Dilarang mengunm
 a. Pengutipan hanya
 b. Pengutipan tidak n

5. Tampilan Menu Nilai K dan E

Tampilan ini merupakan tampilan untuk menentukan nilai K tetangga dan nilai E Exponen pada data judul tugas akhir yang akan diklasifikasikan menggunakan metode NWKNN. Tampilan nilai K dan E dapat dilihat pada gambar berikut ini :

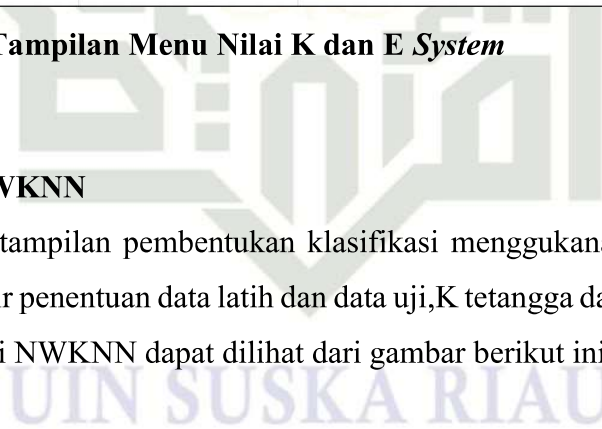


Gambar 5.5 Tampilan Menu Nilai K dan E System

2. Dilarang mengunm
 6. Tampilan Klasifikasi NWKNN
 1. Dilarang mengunm
 a. Pengutipan hanya
 b. Pengutipan tidak n

6. Tampilan Klasifikasi NWKNN

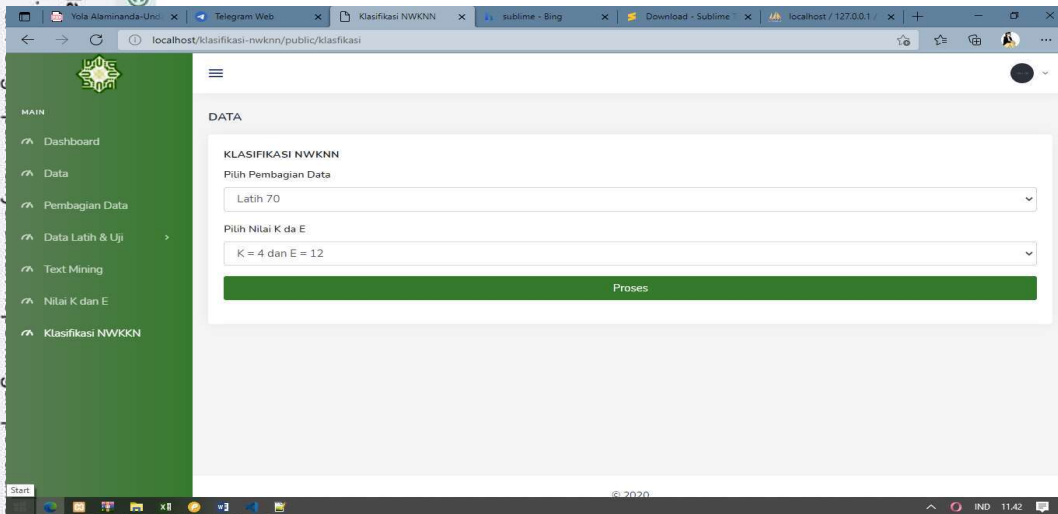
Tampilan ini merupakan tampilan pembentukan klasifikasi menggunakan NWKNN dari data judul Tugas Akhir penentuan data latih dan data uji, K tetangga dan exponen. Tampilan menu klasifikasi NWKNN dapat dilihat dari gambar berikut ini :





2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. H
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.



Gambar 5.6 Tampilan Klasifikasi NWKNN System

Tampilan Proses Klasifikasi NWKNN

Tampilan ini merupakan tampilan dari penentuan data latih dan penentuan pilihan Nilai K tetangga dan e *exponent*. Pada tahap ini merupakan tahap perhitungan akurasi dari klasifikasi data data judul Tugas Akhir yang akan diklasifikasikan menggunakan metode NWKKN. Tampilan proses dari klasifikasi NWKKN ini dapat dilihat dari gambar berikut ini :



2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. H

b. Pengumpulan tidak merugikan kepentingan yang wajar UIN Suska Riau.

The screenshot displays the 'HASIL KLASIFIKASI' (Classification Results) page of the NWKNN system. It shows an overall accuracy of 72.941%, with a value of K=4 and E=12. The data is split into 70% training and 30% testing. A table lists five items with their respective classification metrics and results.

No	Judul	CS	SE	TI	SI	CE	Hasil Klasifikasi	Label	Hasil
1	APLIKASI PEMESANAN AMBULAN KOTA PEKANBARU BERBASIS ANDROID	2.457	0	0	0	1.046	CS	SE	False
2	APLIKASI ANDROID UNTUK PENGHITUNGAN SUARA PILPRES MENGGUNAKAN METODE REAL COUNT DENGAN SMS (SHORT MESSAGE SERVICE)	1.566	1.812	0	0	0	SE	SE	True
3	APLIKASI MONITORING PENGGANTIAN KWH METER BERBASIS ANDROID PADA PT. USAHA PERDANA	3.54	0	0	0	0	CS	SE	False
4	Aplikasi Mengetol Kinerja Pegawai, Studi Kasus: UIN SUSKA RIAU	2.498	0	0	0	1.064	CS	SE	False
5	PENERAPAN METODE ANT COLONY OPTIMIZATION (ACO) UNTUK OPTIMASI ROUTING PADA JARINGAN KOMPUTER	3.54	0	0	0	0	CS	SE	False

Gambar 5.7 Tampilan Proses Klasifikasi NWKNN System

Pengujian

apapun pengujian dilakukan untuk menguji aplikasi yang telah dibuat. Pengujian dilakukan apakah aplikasi yang dibangun sudah memenuhi kebutuhan dan dapat berjalan sesuai dengan tujuan dari hasil yang diinginkan.

ada pengujian ini dilakukan pembagian data latih dan data uji.

1. Pengujian pembagian data latih 70% dan 30% data uji dan nilai K= 4 dan E=6

didapatkan hasil akurasi 75,397%

2. Pengujian pembagian data latih 50% dan 50% data uji dengan nilai K=4 dan E=6

didapatkan hasil akurasi 76,214%

3. Pengujian pembagian data latih 70% dan 30% data uji dengan nilai k=3 dan E=7

didapatkan hasil akurasi 75,397%

4. Pengujian pembagian data latih 50% dan 50% data uji dengan nilai K=3 dan E=7

didapatkan hasil akurasi 76.214%



5.4 Kesimpulan pengujian

Berdasarkan pengujian yang telah dilakukan sebelumnya dapat disimpulkan dari data data pada penelitian ini adalah 412 *record* data mahasiswa Teknik informatika UIN SUSKA riau yang diolah menggunakan aplikasi algoritma metode NWKNN dapat disimpulkan bahwa :

- a. Menghasilkan klasifikasi *text mining* menggunakan algoritma NWKNN dengan 412 *record* data.
- b. Berdasarkan pengujian yang dilakukan menggunakan metode NWKNN didapatkan hasil akurasi lebih dari 70% dengan pembagian data 2/3 dari keseluruhan data serta pembagian dengan 1/2 dari keseluruhan data untuk data latih dan data uji.

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



3.5 Pengujian fungsional

Pada pengujian fungsional ini peneliti menggunakan *whitebox* sebagai pengujian. Hasil dari pengujian *whitebox* adalah sebagai berikut :

Tabel 1. Pengujian *whitebox*

PROSES	CODE	HASIL YANG DIHARAPKAN	HASIL SISTEM	KONDISI
CASE FOLDING	<pre> public function preprocessing(\$datatraining) { //tolowercase dan penghapusan delimiter // \$normalizedText = new \Sastrawi\Filter\TextNormalizer::normalizeText(\$); \$normalizer=new \Sastrawi\Stemmer\Filter\TextNormalizer(); \$output=array(); \$i=0; foreach (\$datatraining as \$dt){ \$kw[\$i]=\$dt->judul; \$i++; } \$i=0; foreach (\$kw as \$k){ \$output[\$i]=\$normalizer->normalizeText(\$k); \$i++; } return \$output; } </pre>	Seluruh data judul Tugas Akhir berubah menjadi huruf kecil/lowercase.	Seluruh data Tugas Akhir berubah menjadi huruf kecil/lowercase	valid

Dilarang mengutip, menyalin, atau menjiplak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

<p>2</p> <p>Hak Cipta Dilindungi Undang-Undang</p> <p>1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber: a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah. b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.</p> <p>2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.</p>	<p>TOKENIZIN</p> <p>Hak cipta milik UIN Suska Riau</p>	<pre> public function frequencygram(\$ngram){ \$y = new \Ngram\Tool\Ngram\Frequency(); \$z=\$y->get(\$ngram); \$output=array(); \$i=0; for (\$i=0;\$i<count(\$z);\$i++){ \$output[\$i][0]=k ey(\$z); \$output[\$i][1]=\$ z[key(\$z)]; next(\$z); } return \$output; </pre>	<p>Judul Tugas Akhir terbentuk per-kata sesuai dengan N-gram yang ditentukan N=1 (Unigram)</p>	<p>Judul Tugas Akhir terbentuk per-kata sesuai dengan N-gram yang ditentukan . N=1 (Unigram)</p>	<p>Valid</p>
	<p>FILTERING</p> <p>State Islamic University of Sultan Syarif Kasim Riau</p>	<pre> public function stopwordremoval(\$datatraining){ \$stopwordFactory = new \Sastrawi\StopWordRemover\St opWordRemoverFactory(); \$stopword=\$stopwordF actory- >createStopWordRemover(); \$output=array(); \$i=0; foreach (\$datatraining as \$dt){ \$output[\$i]=\$sto pword->remove(\$dt); if(\$output[\$i]== ''){ \$output[\$i]= ' '; } \$i++; } return \$output; } </pre>	<p>Menampilkan data Tugas Akhir yang sudah dihapus kata tidak penting yang diambil dari library sastrawi.</p>	<p>Menampilkan data Tugas Akhir yang sudah dihapus kata tidak penting yang diambil dari library sastrawi.</p>	<p>valid</p>

<p>4</p> <p>Hak Cipta Dilindungi Undang-Undang</p> <p>1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber: a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah. b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.</p> <p>2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.</p>	<p>STEAMMIN</p> <p>Hak cipta milik UIN Suska Riau</p>	<pre> public function stemming(\$datatraining){ \$stemmerFactory = new \Sastrawi\Stemmer\StemmerFac tory(); \$stemmer = \$stemmerFactory- >createStemmer(); \$output=array(); // stem \$i=0; foreach (\$datatraining as \$dt){ \$kw[\$i]=\$dt; \$i++; } \$i=0; foreach (\$kw as \$k){ \$output[\$i]=\$ste mmmer->stem(\$k); \$i++; } return(\$output); } </pre>	<p>Menampil kan judul Tugas Akhir yang sudah dihapus awalan dan akhiran yang diambil dari library sastrawi.</p>	<p>Menampil kan judul Tugas Akhir yang sudah dihapus awalan dan akhiran yang diambil dari library sastrawi.</p>	<p>valid</p>
---	---	---	---	---	--------------