



IDENTIFIKASI WORD-BOUNDARY PADA HASHTAG TWEET BAHASA INDONESIA MENGGUNAKAN METODE RULE BASED

TUGAS AKHIR

**Diajukan Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik Pada
Jurusan Teknik Informatika**

Oleh:

ULFA NADA DIANA

11551204660



- 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.

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU

PEKANBARU

2022

LEMBAR PERSETUJUAN

IDENTIFIKASI WORD-BOUNDARY PADA HASHTAG TWEET BAHASA INDONESIA MENGGUNAKAN METODE RULE BASED

TUGAS AKHIR

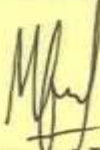
Oleh

ULFA NADA DIANA

NIM. 11551204660


Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir di
Pekanbaru, pada tanggal 7 Juli 2022

Pembimbing I,



Muhammad Fikry, ST, M.Sc
NIP. 19801018 200710 1 002

Pembimbing II,



Yusra, ST, MT
NIP. 19840123 201503 2 001

LEMBAR PENGESAHAN

IDENTIFIKASI WORD-BOUNDARY PADA HASHTAG TWEET BAHASA INDONESIA MENGGUNAKAN METODE RULE BASED

Oleh

ULFA NADA DIANA

NIM. 11551204660

Telah dipertahankan di depan sidang dewan penguji sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik pada Universitas Islam Negeri Sultan Syarif Kasim Riau

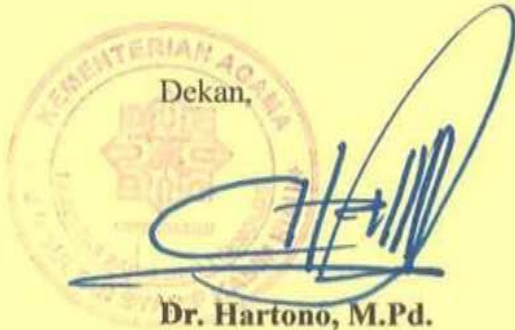
Pekanbaru, 07 Juli 2022

Mengesahkan,

Ketua Jurusan,

Iwan Iskandar, MT.

NIP. 19821216 201503 1 003



Dekan,

Dr. Hartono, M.Pd.

NIP. 19640301 199203 1 003

DEWAN PENGUJI

Ketua : Iwan Iskandar, MT.
Pembimbing I : Muhammad Fikry, S.T., M.Sc.
Pembimbing II : Yusra, S.T., M.T.
Penguji I : Reski Mai Candra, S.T., M.Sc.
Penguji II : Fadhilah Syafria, S.T., M.Kom.



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

Lampiran Surat :
Nomor : Nomor 25/2021
Tanggal : 10 September 2021

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini:

Nama : ULFA NADA DIANA
NIM : 11551204660
Tempat/Tgl. Lahir : Bangkinang / 26 Desember 1994
Fakultas/Pascasarjana : Sains dan Teknologi
Prodi : Teknik Informatika
Judul Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya*:
IDENTIFIKASI WORD-BOUNDARY PADA HASHTAG TWEET
BAHASA INDONESIA MENGGUNAKAN METODE
RULE BASED

Menyatakan dengan sebenar-benarnya bahwa:

1. Penulisan Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* saya ini, saya nyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan Disertasi/Thesis/Skripsi(Karya Ilmiah lainnya)* saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 22 Juli 2021
Yang membuat pernyataan



ULFA NADA DIANA
NIM: 11551204660

*jika sudah ada semua judul karya tulis

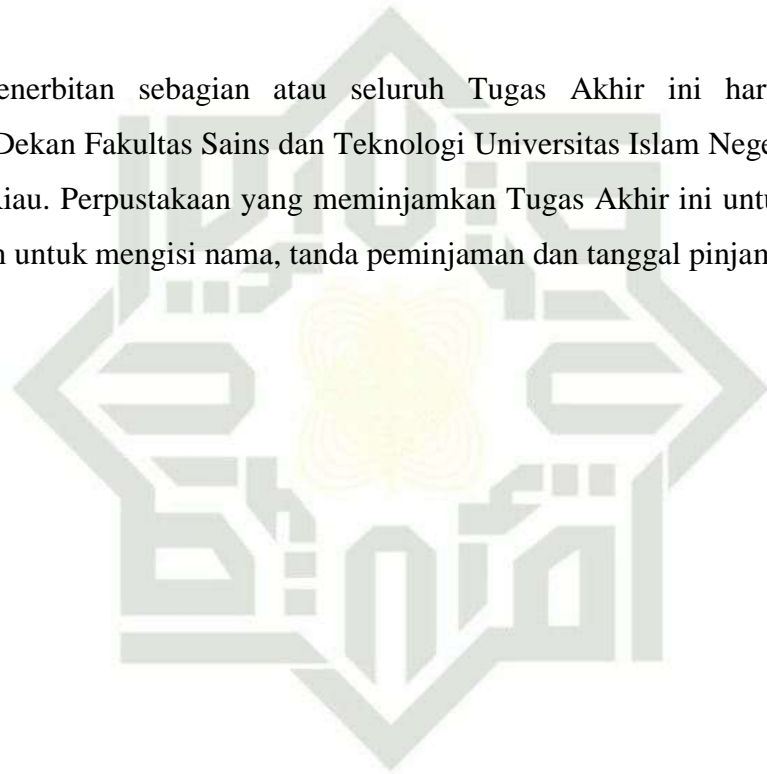


LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dan atas ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan dipergunakan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan dengan izin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggunaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

1. Dianggap sebagai Uraian atau uraian yang tidak dapat dipisahkan dari keseluruhan uraian yang bersangkutan, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
2. Dianggap sebagai Uraian atau uraian yang tidak dapat dipisahkan dari keseluruhan uraian yang bersangkutan, penulisan karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.





LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan didalam daftar pustaka.

Pekanbaru, 7 Juli 2022

Yang membuat pernyataan,

ULFA NADA DIANA

NIM. 11551204660

UIN SUSKA RIAU

- Jakarta: Pustaka Mandiri, 2009.
1. Dikembangkan menjadi bagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Sangkapan dan/atau 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 mengundurkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Hak Cipta dilindungi Undang-undang
1. Dilarang memperjualbelikan atau seluruh karya tulis 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 diperkenankan untuk tujuan yang lain yang melanggar Undang-undang tentang Kekuasaan Yang Wajar UIN Suska Riau.
4. Dilarang menggunakan dan memperjualbelikan dan memperbanyak seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Alhamdulillah Rabbil'alamin. Segala Puji bagi engkau Yaa Allah, Tuhan semesta alam. Puji Syukur ananda ucapkan kepada Allah Subhanahu wa Ta'ala, atas berkat dan pertolongan dari-Nya, diri ini dapat menyelesaikan kewajiban terakhir di dunia perkuliahan, yakni skripsi tugas akhir. Sungguh nikmat yang tak terbayangkan dari engkau Yaa Allah, yang telah memberikan kesempatan bagi diri ini untuk dapat menyelesaikan skripsi tugas akhir.

Skripsi tugas akhir ananda persembahkan untuk orang tua tercinta. Berkat kerja keras dan keringat beserta doa yang selalu orang tua berikan, sehingga ananda dapat menyelesaikan skripsi tugas akhir ini.

Jak lupa pula untuk seluruh teman-teman TIF G'15 dan teman-teman seperjuangan Teknik Informatika angkatan 15 yang telah bersedia membantu dan mendukung ananda selama perkuliahan. Semoga Allah membalas tiap amal kebaikan yang telah kalian berikan dan Allah memberikan kesuksesan untuk kita nantinya setelah menyelesaikan perkuliahan di Teknik Informatika. Aamiin yaa Rabbal'alamin.

ABSTRAK

Bahasa merupakan alat untuk berkomunikasi dalam menyampaikan pikiran, gagasan, konsep dan perasaan. Kesalahan dalam penggunaan bahasa di media sosial menyebabkan penyampaian informasi yang tidak jelas atau ambigu. Terdapat dua jenis kesalahan dalam penulisan, yaitu kesalahan penulisan yang tidak disengaja seperti tipografi dan kesalahan penulisan yang disengaja seperti penulisan kata yang menyatu pada tulisan yang memiliki tagar/*hashtag*. Pada aturan tokenisasi pemisahan kata hanya dilakukan pada teks yang memiliki spasi, enter, tabulasi dan tanda baca lainnya, namun kata yang menyatu pada *hashtag* tidak terdapat tanda baca apapun yang dapat dilakukan pemisahan menggunakan tokenisasi. Oleh karena itu diperlukan adanya identifikasi *word-boundary* untuk memisahkan kata yang menyatu pada *hashtag* agar didapatkan informasi penulisan yang benar. Pemisahan kata dilakukan menggunakan algoritma *rule based* dengan membuat beberapa aturan untuk 10 jenis bentuk penulisan *hashtag*, agar dapat memproses identifikasi *word-boundary* dan mendapatkan hasil pemenggalan kata yang benar. Berdasarkan algoritma dan *rule based* yang telah dirancang, didapatkan akurasi keberhasilan sistem sebesar 98,2% dengan jumlah kesalahan sebesar 9 dari 500 data uji *hashtag*.

Kata Kunci : *Word-Boundary, Bahasa Indonesia, Hashtag, Rule Based*



ABSTRACT

Language is a tool to communicate in conveying thoughts, ideas, concepts and feelings. Errors in the use of language on social media lead to the delivery of unclear or ambiguous information. There are two types of writing errors, namely unintentional writing errors such as typography and intentional writing errors such as writing words that are integrated into posts that have hashtags/hashtags. In the tokenization rules, word separation is only carried out on text that has spaces, enter, tabulations and other punctuation marks, but words that are attached to the hashtag do not contain any punctuation marks that can be separated using tokenization. Therefore, it is necessary to identify a word-boundary to separate the words that are attached to the hashtag in order to obtain the correct writing information. Word separation is done using a rule-based algorithm by making several rules for 10 types of hashtag writing forms, in order to process word-boundary identification and get the correct word-cutting results. Based on the algorithm and rule based that have been designed, the accuracy of the system success is 98.2% with a total error of 9 out of 500 hashtag test data.

Keywords : *Word-Boundary, Indonesian, Hashtag, Rule Based*





KATA PENGANTAR

Assalamu 'alaikum wa rohmatullohi wa barokatuh.

Ahamdillillahi robbil'amin, tak henti-hentinya kami ucapkan kehadiran Allah *Subhanahu wa ta'ala*, yang dengan rahmat dan hidayah-Nya kami mampu menyelesaikan Tugas Akhir dengan judul **“Identifikasi Word-Boundary pada Hashtag Tweet Bahasa Indonesia Menggunakan Metode Rule Based”** ini dengan baik. Tidak lupa bershalawat kepada Nabi dan Rasul-Nya, Nabi Muhammad *Sholahuhu alaihi wa salam*, yang telah membimbing kita sebagai umatnya dari kegelapan hingga menuju jalan kebaikan seperti saat sekarang ini.

Tugas Akhir ini disusun sebagai salah satu syarat untuk mendapatkan gelar sarjana pada Jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau. Selama proses menyelesaikan penelitian laporan tugas akhir ini, penulis mendapatkan banyak pengetahuan, pengalaman, arahan, dan juga dukungan semangat dari banyak pihak yang telah membantu penulis dalam penyusunan laporan ini, baik berupa bantuan materi ataupun berupa motivasi dan dukungan kepada kami. Semua itu tentu terlalu banyak bagi kami untuk membalasnya, namun pada kesempatan ini penulis hanya dapat mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Khairunnas Rajab, M.Ag., selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
3. Bapak Iwan Iskandar, ST., MT selaku Ketua Prodi Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Bapak Reski Mai Candra, ST., M.Sc selaku Sekretaris Prodi Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.



5. Bapak Novriyanto,ST.,M.Sc selaku dosen pembimbing akademik yang telah memberi arahan dan saran kepada penulis selama kuliah di jurusan Teknik Informatika.

Bapak Muhammad Fikry,ST.,M.Sc dan Ibu Yusra,M.T selaku dosen pembimbing Tugas Akhir penulis yang telah sangat banyak meluangkan waktu, ilmu, arahan, bimbingan serta kesabaran yang sangat banyak dalam membimbing penulis untuk menyelesaikan Tugas Akhir ini.

Bapak Reski Mai Candra,ST.,M.Sc selaku penguji I yang telah memberikan arahan, kritik, dan saran kepada penulis dalam penyusunan Tugas Akhir ini.

Ibu Fadhilah Syafria ST, M.Kom selaku penguji II yang telah memberikan arahan, kritik, dan saran kepada penulis dalam penyusunan Tugas Akhir ini.

Bapak dan Ibu Dosen Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau yang telah memberikan ilmunya kepada penulis selama perkuliahan.

Kepada orang tua tercinta, Ayahanda H.Nazwir Z dan Ibunda Hj.Asmarida, serta kakak penulis Rafika Nada Rizki, S.Pd dan Yesi Nada Sari, SKM yang selalu senantiasa mendoakan, dan memberikan semangat selama penulis kuliah di jurusan Teknik informatika serta selalu mengingatkan penulis agar dapat menyelesaikan Tugas Akhir ini.

1. Kepada teman-teman seperjuangan TIF G angkatan 2015 yang selalu mendoakan dan memberi dukungan serta semangat kepada penulis.

2. Kepada teman-teman seangkatan TIF 2015, para senior TIF dan junior TIF yang tidak bisa penulis sebutkan satu persatu namanya.

3. Semua pihak yang terlibat langsung maupun tidak langsung dalam pelaksanaan Tugas Akhir ini yang tidak dapat penulis sebutkan satu persatu.

4. Seluruh pihak yang belum penulis cantumkan, terima kasih atas dukungannya, baik material maupun spiritual.

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.



Hak Cipta Dilindungi Undang-Undang

1. Di larang menyalin sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Paragraf, hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Paragraf tidak mengizinkan kepentingan yang wajar UIN Suska Riau.
 2. Di larang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak cipta milik UIN Suska Riau

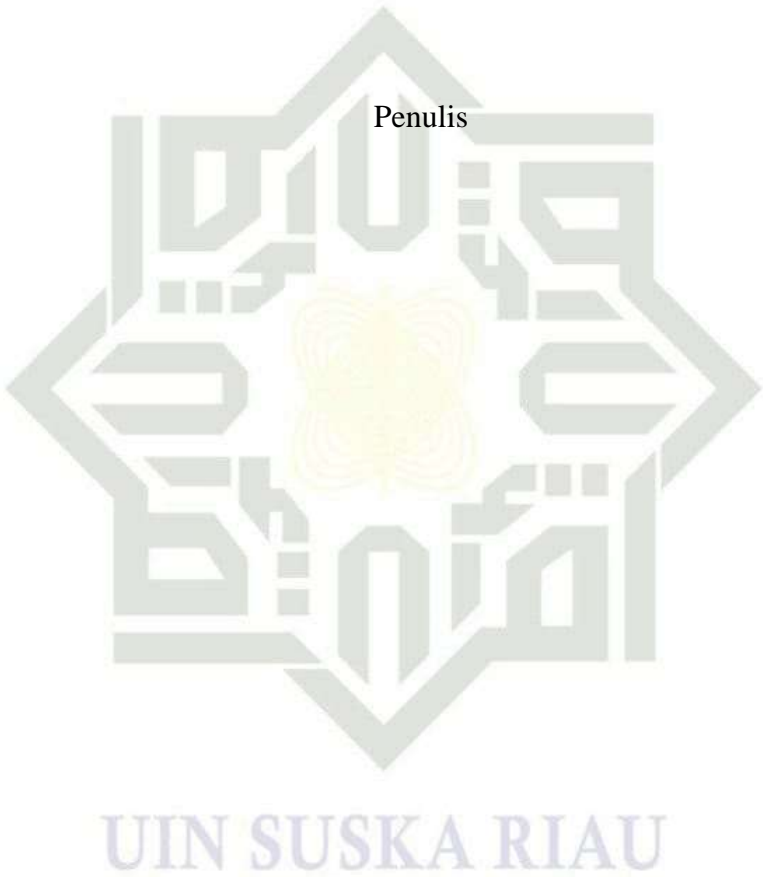
State Islamic University of Sultan Syarif Kasim Riau

Penulis menyadari bahwa dalam penulisan laporan ini masih banyak kesalahan dan kekurangan, oleh karena itu kritik dan saran yang sifatnya membangun sangat kami harapkan untuk kesempurnaan laporan ini. Akhirnya penulis berharap semoga laporan kripsi ini dapat memberikan sesuatu yang bermanfaat bagi siapa saja yang membacanya.

Wassalamu'alaikum wa rohmatullohi wa barokatuh.

Pekanbaru, 7 Juli 2022

Penulis





DAFTAR ISI

1.	LEMBAR PERSETUJUAN.....	i
2.	LEMBAR PENGESAHAN	ii
3.	LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL	iii
4.	LEMBAR PERNYATAAN.....	v
5.	LEMBAR PERSEMBAHAN.....	vi
6.	ABSTRAK	vii
7.	ABSTRACT	viii
8.	KATA PENGANTAR	ix
9.	DAFTAR ISI.....	xii
10.	DAFTAR GAMBAR	xv
11.	DAFTAR TABEL.....	xvi
12.	DAFTAR SIMBOL	xvii
13.	BAB I PENDAHULUAN	1
14.	1.1. Latar Belakang	1
15.	1.2. Rumusan Masalah	4
16.	1.3. Batasan Masalah.....	4
17.	1.4. Tujuan Penelitian.....	4
18.	1.5. Sistematika Penulisan.....	4
19.	BAB II LANDASAN TEORI.....	6
20.	2.1. <i>Natural Language Processing</i>	6
21.	2.2. Tokenisasi.....	6
22.	2.3. <i>Word Boundary</i>	7
23.	2.4. <i>Rule Based</i>	7
24.	2.5. <i>Twitter</i>	8
25.	2.6. <i>Hashtag (#)</i>	8
26.	2.8. Penelitian Terkait	8
27.	BAB III METODOLOGI PENELITIAN	13
28.	3.1. Identifikasi Masalah	14
29.	3.2. Studi Pustaka	14
30.	3.3. Pengumpulan Data	14

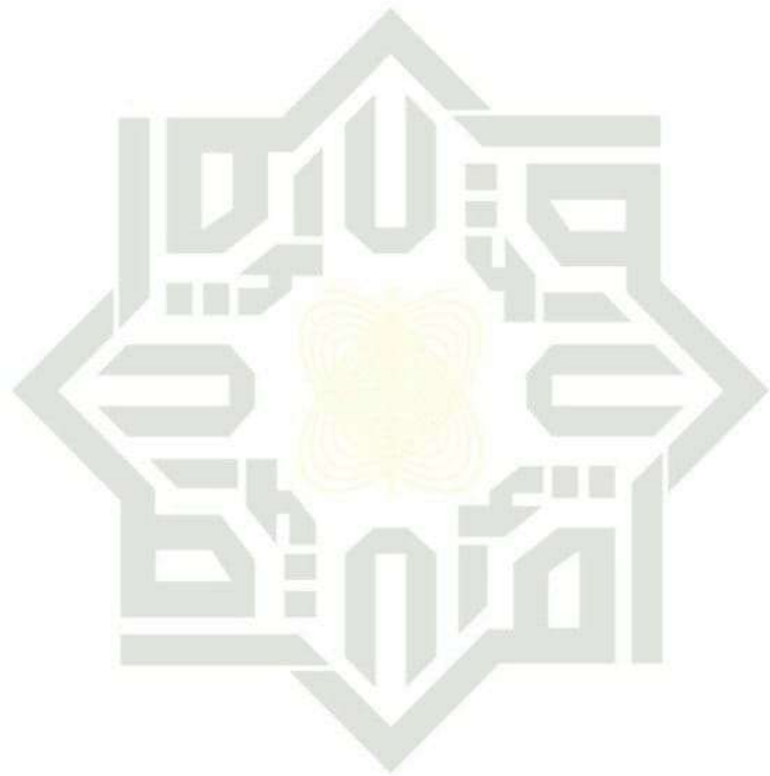
LAMPIRAN A	65
LAMPIRAN B	96
LAMPIRAN C	97



UIN SUSKA RIAU

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau



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

DAFTAR GAMBAR

Gambar 3. 1 Flowchart Metodologi Penelitian	13
Gambar 3. 2 Flowchart Analisa Identifikasi Word-Boundary	16
Gambar 4. 1 Analisa <i>Flowchart</i> Identifikasi <i>Word-Boundary</i>	31
Gambar 4. 2 Pseudocode Import Hashtag	35
Gambar 4. 3 Pseudocode Import Kamus.....	35
Gambar 4. 4 Pseudocode Bentuk Penulisan Huruf Kecil Semua.....	36
Gambar 4. 5 Pseudocode Bentuk Penulisan Huruf Besar Semua	36
Gambar 4. 6 Pseudocode Bentuk Penulisan Huruf Pertama Setiap Kata Huruf Besar .	36
Gambar 4. 7 Pseudocode Bentuk Penulisan Huruf Pertama Setiap Kata Huruf Besar dan Kapital Singkatan	37
Gambar 4. 8 Pseudocode Bentuk Penulisan Huruf Pertama Huruf Besar dan Mengandung Angka	37
Gambar 4. 9 Pseudocode Bentuk Penulisan Huruf Kecil Semua dan Tanda Baca () Sebagai Pemisah Kata	38
Gambar 4. 10 Pseudocode Bentuk Penulisan Huruf Besar Semua dan Tanda Baca () Sebagai Pemisah Kata	38
Gambar 4. 11 Pseudocode Bentuk Penulisan Angka Pengganti Huruf, Huruf Pertama Setiap Kata Huruf Besar.....	39
Gambar 4. 12 <i>Interface</i> Halaman Utama.....	41
Gambar 4. 13 <i>Interface</i> Halaman Data <i>Hashtag</i>	42
Gambar 4. 14 <i>Interface</i> Tes Identifikasi	42
Gambar 4. 15 <i>Interface</i> Hasil Identifikasi.....	43
Gambar 5. 1 Halaman Utama.....	46
Gambar 5. 2 Halaman Data <i>Hashtag</i>	46
Gambar 5. 3 Halaman Data Kamus.....	47
Gambar 5. 4 Halaman Hasil Identifikasi <i>Word-boundary</i>	48

DAFTAR TABEL





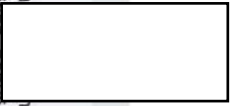


UIN SUSKA RIAU
 Disajikan dengan tujuan untuk memberikan informasi yang akurat dan dapat dipercaya kepada masyarakat luas. Diizinkan untuk digunakan secara terbatas untuk keperluan akademik dan penelitian. Tidak diperbolehkan untuk diperjualbelikan atau untuk tujuan komersial.

Tabel 4.1	Tabel Data <i>Hashtag</i>	20
Tabel 4.2	Jenis Penulisan <i>Hashtag</i>	21
Tabel 4.3	Tabel Data Kamus	23
Tabel 4.4	Tabel Hasil Identifikasi Kata	34
Tabel 4.5	Tabel Periksa Ketersediaan Kata	34
Tabel 4.6	Tabel <i>User</i> Pada <i>Database</i>	39
Tabel 4.7	Tabel <i>Hashtag</i> Pada <i>Database</i>	40
Tabel 4.8	Tabel Hasil pada <i>Database</i>	40
Tabel 5.1	Contoh tabel cek ketersediaan kata pada kamus	49
Tabel 5.2	Tabel nilai ketersediaan kata pada kamus	49
Tabel 5.3	Tabel Perbandingan Hasil Sistem dengan Validator	51
Tabel 5.4	Tabel Akurasi Berdasarkan Jenis Penulisan <i>Hashtag</i>	58
Tabel 5.5	Tabel pengujian halaman <i>login</i>	59
Tabel 5.6	Tabel pengujian halaman beranda	59
Tabel 5.7	Tabel pengujian halaman data <i>hashtag</i>	60
Tabel 5.8	Tabel pengujian halaman data kamus	60
Tabel 5.9	Tabel pengujian halaman hasil identifikasi	60

State Islamic University of Sultan Syarif Kasim Riau

UIN SUSKA RIAU

DAFTAR SIMBOL

Chart	Keterangan
	<p><i>Terminator</i> : Simbol <i>terminator</i> (Mulai/Selesai) merupakan tanda bahwa sistem akan dijalankan atau berakhir.</p>
	<p>Data : Simbol yang digunakan untuk mendeskripsikan data yang digunakan.</p>
	<p>Proses : Simbol yang digunakan untuk melakukan pemrosesan data baik oleh <i>user</i> maupun komputer (sistem).</p>
	<p>Verifikasi : Simbol yang digunakan untuk memutuskan apakah valid atau tidak validnya suatu kejadian.</p>
	<p>Arus Data: Simbol yang digunakan untuk menggambarkan arus data di dalam sistem.</p>



BAB I PENDAHULUAN

1.1. Latar Belakang

Bahasa merupakan sumber awal manusia memperoleh pemahaman dan ilmu pengetahuan untuk memahami apa yang ada disekitarnya sehingga memiliki ilmu pengetahuan dan keahlian. Secara sederhana, bahasa dapat diartikan sebagai alat untuk berinteraksi dalam menyampaikan pikiran, gagasan, konsep atau perasaan. (Murti, 2015). Bahasa adalah alat komunikasi lingual manusia baik lisan maupun tulisan. Peran Bahasa Indonesia yang baik dan benar sebagai bahasa resmi nasional memiliki arti yang sangat penting. (Adriyani, Santiyasa, & Muliantara, 2012).

Bahasa Indonesia merupakan bahasa nasional dan bahasa resmi di Indonesia. Bahasa Indonesia adalah alat komunikasi yang dipergunakan oleh masyarakat Indonesia untuk keperluan sehari-hari. Pemakaian bahasa Indonesia dalam konteks bahasa nasional dapat digunakan dengan bebas menggunakan ujaran baik lisan maupun tulisan. Bahasa Indonesia mampu mempersatukan komunikasi bermacam-macam kelompok etnis dan ras di Indonesia (Fallis, 2013)

Text mining merupakan proses menambang data yang berupa teks dimana sumber data biasanya didapatkan dari teks dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar kata. Tujuan dari *text mining* adalah mengekstrak informasi yang berguna dari sumber data. Jadi, sumber data yang digunakan pada *text mining* adalah sekumpulan dokumen atau teks yang memiliki format yang tidak terstruktur melalui identifikasi dan eksplorasi pola yang menarik. (Soepomo, 2014).

Tokenisasi merupakan suatu proses pemotongan atau pemisahan kalimat menjadi beberapa bagian atau kata berdasarkan karakter spasi dan penghapusan karakter tertentu seperti tanda baca pada waktu yang bersamaan. Pemecahan kalimat menjadi kata-kata tunggal dilakukan dengan cara *scan* (memindai) kalimat sebagai pemisah (*Delimiter*) berupa *whitespace* (spasi, tab dan *new line*). (Astuti, Rachmat C., & Lukito, 2017).

UN 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, penerbitan atau terjemahan.
b. Pengutipan tidak mengizinkan kepentingan yang bersifat komersial.
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UN Suska Riau.

Word-boundary adalah awalan dan akhiran dari sebuah kata. Dalam penulisan batas kata (*word-boundary*) secara konvensional ditunjukkan dengan adanya spasi antar kata. Kesalahan *word-boundary* adalah kesalahan kata dimana kata yang salah diakibatkan oleh penempatan spasi yang tidak sesuai. Kesalahan *word-boundary* ditangani dengan mencari dua kata pada Kamus Besar Bahasa Indonesia (KBBI), apabila dua kata tersebut digabungkan akan sama dengan kata yang dianggap salah eja *word-boundary*(Rohim & Purnamasari, 2016). *Word-boundary* merupakan perangkat lunak yang dirancang untuk melakukan pemeriksaan kebenaran ejaan kata yang benar(Habibi, 2009).

Media sosial telah melahirkan bentuk-bentuk baru dalam berkomunikasi dan berinteraksi dengan masyarakat. Media sosial kini sudah menjadi gaya hidup masyarakat. Di media sosial, tagar dikenal dengan istilah *hashtag*. Penggunaan *hashtag* ini bermanfaat untuk mengelompokkan pesan yang masuk agar dapat dicari dengan mudah. Sebuah gerakan bisa terkenal karena *hashtag* (#) yang unik atau kontroversial. (Mustofa, 2019).

Hashtag merupakan sekumpulan karakter alfanumerik yang didahului dengan simbol '#', *hashtag* menjadi populer dan menyebar ke platform media sosial lainnya. *Hashtag* disebut juga dengan kata kunci untuk mencari informasi. Dengan *hashtag* kita dapat menyaring sebuah informasi karena tujuan awal sebuah kata bertanda pagar (#) memang digunakan untuk mengelompokkan sebuah topik(Fallis, 2013)

Mendeteksi batas token dan kalimat adalah langkah penting dalam tahap *pre-processing* di aplikasi *natural language processing* karena sebagian besar beroperasi baik pada tingkat kata atau kalimat (penandaan paruh waktu, *parsing*, terjemahan mesin). Kesalahan dalam penggunaan bahasa di sosial media menyebabkan penyampaian informasi yang tidak jelas atau ambigu, oleh karna itu diperlukan tokenisasi untuk menanganinya. Tantangan utama pada tugas tokenisasi yaitu adanya ambiguitas karakter tertentu dalam alphabet dan dari tidak adanya batas kata yang diberi dalam penulisan simbolis (Jurish, 2013).

Terdapat dua jenis kesalahan dalam penulisan suatu kata ataupun kalimat, yaitu kesalahan penulisan yang tidak disengaja seperti tipografi dan kesalahan

penulisan yang disengaja. Sebagai contoh yaitu pada teks yang memiliki *hashtag*, seperti ‘Selamat hari selasa #harisantrinasional’, teks yang memiliki *hashtag* tersebut seharusnya dapat dipisah menjadi ‘hari santri nasional’, kesalahan seperti

ini merupakan kesalahan penulisan yang disengaja. Kata yang menyatu di dalam *hashtag* tersebut tidak bisa dipisah menggunakan tokenisasi. Hal ini menyebabkan kesalahan dalam ejaan sehingga menyebabkan arti yang berbeda. Kesalahan tersebut juga dapat membuat mesin sulit mendeteksi kata dari sebuah teks. Diperlukan adanya proses tokenisasi ulang menggunakan *word-boundary* pada kalimat tersebut untuk memisahkan teks yang tidak memiliki pemisah kata untuk mendapatkan makna sebenarnya.

Penelitian yang sudah dilakukan sebelumnya mengenai tokenisasi yaitu penelitian (Jurish, 2013) yaitu *Word and Sentence Tokenization with Hidden Markov Models* menyajikan metode baru untuk memperkirakan kalimat dan kata pada token *running text* dengan koping tahap *scanner* berbasis aturan panjang lebar dengan *Hidden Markov Model* lebih fitur bundel scan-segmen menggunakan fitur biner tersembunyi busur, bos, dan eos untuk mewakili kehadiran atau tidak adanya batas-batas yang sesuai.

Penelitian (Huang, Simon, Hsieh, & Pr, 2007) yaitu *Rethinking Chinese Word Segmentation: Tokenization, Character, Classification, or Wordbreak Identification* berhasil memperkenalkan konsep klasifikasi Segmentasi kata Cina dari serangkaian batasan karakter ke salah satu batas kata dan batas-batas *non-kata*.

Penelitian (Koehn & Knight, 2003) yaitu *Empirical Methods for Compound Splitting* menghasilkan Aplikasi NLP seperti *Machine Translation* (MT) dengan memperkenalkan metode untuk belajar aturan pemisahan dari satu bahasa dan korpora paralel. Hasil menunjukkan akurasi 99,1% dan keuntungan kinerja untuk MT dari 0,39 BLEU pada bahasa Jerman-Inggris tugas terjemahan frase kata benda.

Sehingga dari permasalahan diatas maka penulis ingin memecahkan masalah tersebut dengan melakukan penelitian dengan judul “Identifikasi *word-boundary* pada *hashtag tweet* Bahasa Indonesia” yang akan dilakukan

menggunakan metode *Rule Based*, dan kombinasi *Sliding Window* bersama *Dictionary Lookup* dan *Approximate String Matching*.

1.2. Rumusan Masalah

Berdasarkan latar belakang masalah diatas, yang menjadi pokok permasalahan dalam penelitian ini adalah “Bagaimana membuat algoritma Identifikasi *word-boundary* pada *hashtag tweet* Bahasa Indonesia dan menguji akurasi algoritma *word-boundary* Bahasa Indonesia Menggunakan *Rule Based*”.

1.3. Batasan Masalah

Agar penyusunan penelitian Tugas Akhir ini tidak keluar dari pokok permasalahan yang telah dirumuskan, maka diberikan beberapa batasan masalah antara lain:

1. Data yang digunakan adalah data yang berasal dari sosial media yaitu *Twitter* dari berbagai akun dalam bahasa Indonesia.
2. Data yang diambil merupakan data *tweet* yang memiliki tagar/*hashtag*.
3. Jumlah data yang digunakan sebanyak 500 data *tweet*.
4. Identifikasi *word-boundary* menggunakan metode *Rule Based*.

1.4. Tujuan Penelitian

- Adapun tujuan dilakukannya penelitian Tugas Akhir ini adalah:
1. Membuat algoritma *word-boundary* bahasa Indonesia.
 2. Menguji akurasi algoritma *word-boundary* bahasa Indonesia.

1.5. Sistematika Penulisan

Sistematika penulisan adalah gambaran umum untuk menyusun laporan skripsi ini, disusun oleh penulis yang didalamnya terdiri dari enam bab dan masing masing bab akan dibahas sebagai berikut:

BAB I PENDAHULUAN

Berisikan tentang latar belakang permasalahan, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.



- BAB I**
- BAB II**
- BAB III**
- BAB IV**
- BAB V**
- BAB VI**
- 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.

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

LANDASAN TEORI

Pada bab ini akan dibahas teori-teori yang berhubungan dengan penelitian yang di teliti yaitu Pengertian *Natural Language Processing*, Pengertian Tokenisasi, Pengertian *Word-Boundary*, Pengertian *Twitter*, Pengertian *Application Programming Interface*, dan Penelitian Terkait.

METODOOGI PENELITIAN

Bab ini berisi tentang rangkaian tahapan dalam penelitian, tahapan pengumpulan data, analisa kebutuhan aplikasi, perancangan perangkat lunak, implementasi, pengujian sistem dan waktu penelitian.

ANALISA DAN PERANCANGAN

Bab ini berisikan tentang analisa kebutuhan, kemudian melakukan perancangan terhadap desain yang dibutuhkan untuk merancang dan membangun aplikasi tersebut.

HASIL DAN IMPLEMENTASI

Bab ini berisikan implementasi, bentuk hasil dari sistem yang telah dirancang sebelumnya dan melakukan pengujian.

KESIMPULAN DAN SARAN

Bab ini merupakan kesimpulan dan saran yang berisikan tentang kesimpulan dan saran dari penulis.



BAB II LANDASAN TEORI

2.1. *Natural Language Processing*

Natural Language Processing (NLP) dapat didefinisikan sebagai kemampuan suatu komputer untuk memproses bahasa, baik lisan maupun tulisan yang digunakan oleh manusia dalam percakapan sehari-hari. Untuk proses komputasi, bahasa harus direpresentasikan sebagai rangkaian simbol yang mengikuti aturan tertentu. Secara sederhana, NLP mencoba untuk membuat komputer dapat mengerti perintah-perintah yang ditulis dalam standar bahasa manusia. Terdapat beberapa alasan yang menyulitkan dalam proses NLP, yaitu masalah ambiguitas atau makna ganda dan jumlah kosa kata (*vocabulary*) yang besar dan berkembang dari waktu ke waktu (Lisangan, 2015).

Natural Language Processing adalah pembuatan program yang memiliki kemampuan untuk memahami bahasa manusia. Pada prinsipnya bahasa alami adalah suatu bentuk representasi dari suatu pesan yang ingin dikomunikasikan antar manusia. Tujuan dari NLP adalah melakukan proses pembuatan model komputasi dari bahasa yang digunakan sehingga dapat terjadi suatu interaksi antara manusia dengan komputer dengan NLP (Wangsanegara & Subaeki, 2015).

2.2. *Tokenisasi*

Tokenisasi merupakan suatu proses memisahkan sebuah kata yang terdapat dalam kalimat, paragraf, atau halaman menjadi token atau potongan tunggal yang berdiri sendiri. Tahap ini juga menghilangkan karakter tertentu seperti tanda baca dan mengubah semua token ke huruf kecil (*lower case*) (Robinson, 2014).

Tokenisasi adalah pemotong kalimat berdasarkan tiap kata yang tersusun. Pemecahan kalimat dilakukan dengan menscan kalimat pemisah (delimiter) *whitespace* (spasi, tab, dan new line). Secara umum tokenisasi merupakan tahap memecahkan sekumpulan karakter dalam suatu teks kedalam satu kata. Sekumpulan karakter tersebut dapat berupa karakter *whitespace*, seperti enter, tabulasi, spasi. Namun untuk karakter petik tunggal (,), titik (.), semikolon (;), titik dua (:), atau

lainnya, juga dapat memiliki peran yang cukup banyak sebagai pemisah kata. Sebuah titik (.) biasanya untuk tanda akhir kalimat, tapi dapat juga muncul dalam singkatan, inisial orang, alamat internet, dll. Kemudian tanda *hyphen* (-) biasanya muncul untuk menggabungkan dua token yang berbeda untuk membentuk token tunggal. Tapi dapat pula ditemukan untuk menyatakan rentang nilai, kata berulang, dan sebagainya. Atau karakter *slash* (/) sebagai pemisah file atau direktori atau *url* ataupun untuk menyatakan “dan atau” (Susandi & Sholahudin, 2016).

2.3. Word Boundary

Word-boundary adalah awalan dan akhiran dari sebuah kata. Dalam penulisan batas kata (*word-boundary*) secara konvensional ditunjukkan dengan adanya spasi antar kata. Kesalahan *word-boundary* adalah kesalahan kata dimana kata yang salah diakibatkan oleh penempatan spasi yang tidak sesuai. Kesalahan *word-boundary* ditangani dengan mencari dua kata pada kamus Indonesia (KBBI), apabila dua kata tersebut digabungkan akan sama dengan kata yang dianggap salah pada *word-boundary* (Rohim & Purnamasari, 2016). *Word-boundary* merupakan perangkat lunak yang dirancang untuk melakukan pemeriksaan kebenaran ejaan dalam dokumen untuk membantu mendapatkan ejaan yang benar (Habibi, 2009).

2.4 Rule Based

Sistem berbasis aturan (*rule based*) digunakan pada saat ekstraksi informasi untuk *plotting* data yang akan diproses dan perubahan karakter berdasarkan susunan perubahan dan perubahan karakter berdasarkan aturan tata bahasa. Dimana suatu bahasa tidak saja berkisar pada persoalan bagaimana melambangkan bunyi-bunyi ujaran serta bagaimana menempatkan tanda-tanda baca dan sebagainya, tetapi juga meliputi hal-hal seperti: bagaimana memotong-motong suku kata, bagaimana menggabungkan kata-kata, baik dengan imbuhan-imbuhan maupun antara kata dengan kata. Pemotongan itu harus berguna terutama bagaimana kita harus memisahkan huruf-huruf itu pada akhir suatu baris, bila baris itu tidak memungkinkan kita menuliskan seluruh kata di sana. Kecuali itu, penggunaan huruf kapital juga merupakan unsur penting yang harus diperhatikan dalam penulisan dengan ejaan yang tepat. (Rohim & Purnamasari, 2016).

2.5. Twitter

Twitter adalah sebuah situs web yang menawarkan jaringan sosial berupa mikroblog sehingga memungkinkan penggunanya untuk mengirimkan dan membaca pesan yang disebut “kicauan” (*tweet*), yang bebas mengekspresikan sesuatu seperti keluhan/kritik terhadap kebijakan pemerintah. “Kicauan” adalah berupa teks tulisan hingga 140 karakter yang ditampilkan pada halaman profil penggunanya (Puspita, 2016).


2.6. Hashtag (#)


Penggunaan kode pagar atau *hashtag* (#) sebagai salah satu penanda dari fitur *Twitter* mempermudah sebuah topik untuk dibicarakan. Fitur ini biasanya digunakan untuk membedakan antara sebuah topik dengan topik lain yang memiliki kalimat yang mirip atau serupa di dalam aktivitas di media sosial *Twitter*. Penggunaan tanda pagar (*hashtag*) di *Twitter* pada negara-negara lain tujuannya agar orang lain paham tentang topik yang dimaksudkan oleh pengguna *hashtag*. Namun kenyataan yang lain terjadi di Indonesia, di mana tujuan penggunaan tanda pagar tersebut salah satunya adalah untuk menaikkan popularitas sebuah topik tertentu. Dengan penggunaan *hashtag* ini pengguna lainnya ikut melakukan *retweet* dan juga membuat *hashtag* yang sama. Hal ini untuk meningkatkan intensitas tentang topik yang dimaksud (Juditha, 2015).


2.8. Penelitian Terkait


Penelitian terkait pada pembangunan aplikasi terdapat pada table 2.1 berikut:

No	Peneliti	Judul	Hasil
1.	Bryan Jurish, Kay-Michael Würzner	Word and Sentence Tokenization with Hidden Markov Models	Menyajikan metode baru untuk memperkirakan kalimat dan kata batas token <i>running text</i> dengan kopling tahap <i>scanner</i> berbasis aturan panjang lebar dengan Hidden Markov Model serta fitur bundel scan-segmen menggunakan fitur biner tersembunyi busur, bos, dan eos untuk mewakili kehadiran

			atau tidak adanya batas-batas yang sesuai.
<p>2. Dilarang mengemukakan dan mempe- nyawakan karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.</p>	<p>Chi-Ren Tseng Simon, Kai-Hsiang Hsieh, Arrent Peyot</p>	<p>Rethinking Chinese Word Segmentation: Tokenization, Character Classification, or Wordbreak Identification</p>	<p>Berhasil memperkenalkan konsep klasifikasi Segmentasi kata Cina dari serangkaian batasan karakter ke salah satu batas kata dan batas-batas <i>non</i>-kata.</p>
<p>3. Dilarang mengemukakan dan mempe- nyawakan karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.</p>	<p>Philipp Koehn</p>	<p>Empirical Methods for Compound Splitting</p>	<p>Menghasilkan Aplikasi NLP seperti <i>Machine Translation</i> (MT) dengan memperkenalkan metode untuk belajar aturan pemisahan dari satu bahasa dan korpora paralel. Hasil menunjukkan akurasi 99,1% dan keuntungan kinerja untuk MT dari 0,039 BLEU pada bahasa Jerman-Inggris tugas terjemahan frase kata benda.</p>
<p>4. Dilarang mengemukakan dan mempe- nyawakan karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.</p>	<p>Eva Pettersson, Beáta Megyesi, Joakim Nivre</p>	<p>Normalisation of Historical Text Using Context-Sensitive Weighted Levenshtein Distance and Compound Splitting</p>	<p>Dengan pendekatan yang disajikan, proporsi token dengan ejaan yang identik dengan emas ejaan standar meningkat dari garis dasar 64,6% (untuk teks yang tidak normal) menjadi 77,0%, menggunakan tidak ada data yang dianotasi secara manual untuk pelatihan.</p>
<p>5. Dilarang mengemukakan dan mempe- nyawakan karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.</p>	<p>Ralf D. Brown</p>	<p>Corpus-Driven Splitting of Compound Words</p>	<p>Sebuah metode disajikan untuk memisahkan kata majemuk ke dalam konstituennya berdasarkan kata-kata serumpun</p>

		<p>dalam bahasa lain dari corpus paralel.</p>
<p>6. Jonathan Webster, Hanya Kit</p>	<p>Tokenization as the Initial Phase in NLP</p>	<p>Pengertian kata dan token dibahas dan didefinisikan dari sudut pandang leksikografi dan pragmatis</p> <p>implementasi, masing-masing. Pendekatan praktis untuk identifikasi token majemuk dalam bahasa Inggris, seperti idiom, phrasal kata kerja dan ekspresi tetap, dikembangkan</p>
<p>7. Diki Susandi, Husep Sholahudin</p>	<p>Pemanfaatan <i>Vector Space</i> Model pada Penerapan Algoritma Nazief Adriani, KNN dan Fungsi <i>Similarity Cosine</i> untuk Pembobotan IDF dan WIDF pada <i>Prototipe</i> Sistem Klasifikasi Teks Bahasa Indonesia</p>	<p>Menganalisis efektifitas model sistem klasifikasi / kategorisasi dokumen dalam penerapan <i>vector space model</i></p> <p>berdasarkan pembobotan <i>term</i> dokumen dan <i>query</i>, juga menerapkan metode <i>stemming</i> Bahasa Indonesia dengan algoritma <i>nazief adriani</i>, menghasilkan nilai <i>similarity</i> dengan fungsi <i>cosine</i> yang berpengaruh pada pemeringkatan hasil kategorisasi dokumen yang relevan.</p>

<p>8.</p> 	<p>Gregory Grefenstette Pasi Tapanainen</p> <p>Hak Cipta Dilindungi Undang-Undang 1. Dilarang mengutip sebagian atau seluruh karya atau hasil penelitian, penemuan, atau temuan yang telah dipublikasikan tanpa izin dari lembaga pembuat publikasi. 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.</p>	<p>What is a word, What is a sentence? Problems of Tokenization</p>	<p>Definisi apa yang dianggap sebagai token dapat didefinisikan secara tepat dan jelas. ada banyak cara untuk memutuskan apa yang akan dianggap sebagai unit untuk pendekatan komputasi untuk teks. hadir peran tokenization, metode Tokinizing, tata bahasa untuk mengenali akronim, singkatan, dan ekspresi reguler</p>
	<p>Made Muni Adriyani, Wayan Santiyasa, Agus Muliantara</p>	<p>Implementasi Algoritma <i>LEVENSHTEIN DISTANCE</i> dan metode <i>EMPIRIS</i> untuk menampilkan saran perbaikan kesalahan pengetikan dokumen bahasa Indonesia</p>	<p>Algoritma yang digunakan untuk memberikan saran perbaikan adalah <i>Levenshtein Distance</i> yang dapat menghitung keterkaitan antar string dan menghitung jumlah keterbedaan antar dua string. Disamping itu kesalahan penulisan juga dapat disebabkan kurangnya spasi antar kata sehingga kata tersebut tidak mengandung makna maka dengan menerapkan metode empiris diharapkan dapat menjadi jalan keluarnya.</p>
<p>10.</p>	<p>Ema Urtami, Sri Hartati</p>	<p>Pendekatan Metode Rule Based dalam Mengalihbahasakan Teks Bahasa Inggris ke Teks Bahasa Indonesia</p>	<p>Penerjemah teks bahasa Inggris ke teks bahasa Indonesia ini mampu mengenali kumpulan kalimat dalam beberapa paragraf untuk sekali eksekusi input. Penggunaan metode Rule Based dengan parsing tree dan aturan produksi dalam penerjemah teks bahasa Inggris ke teks bahasa Indonesia dapat menangani perbedaan aturan tata bahasa dari kedua bahasa tersebut sehingga menghasilkan terjemahan yang sesuai.</p>

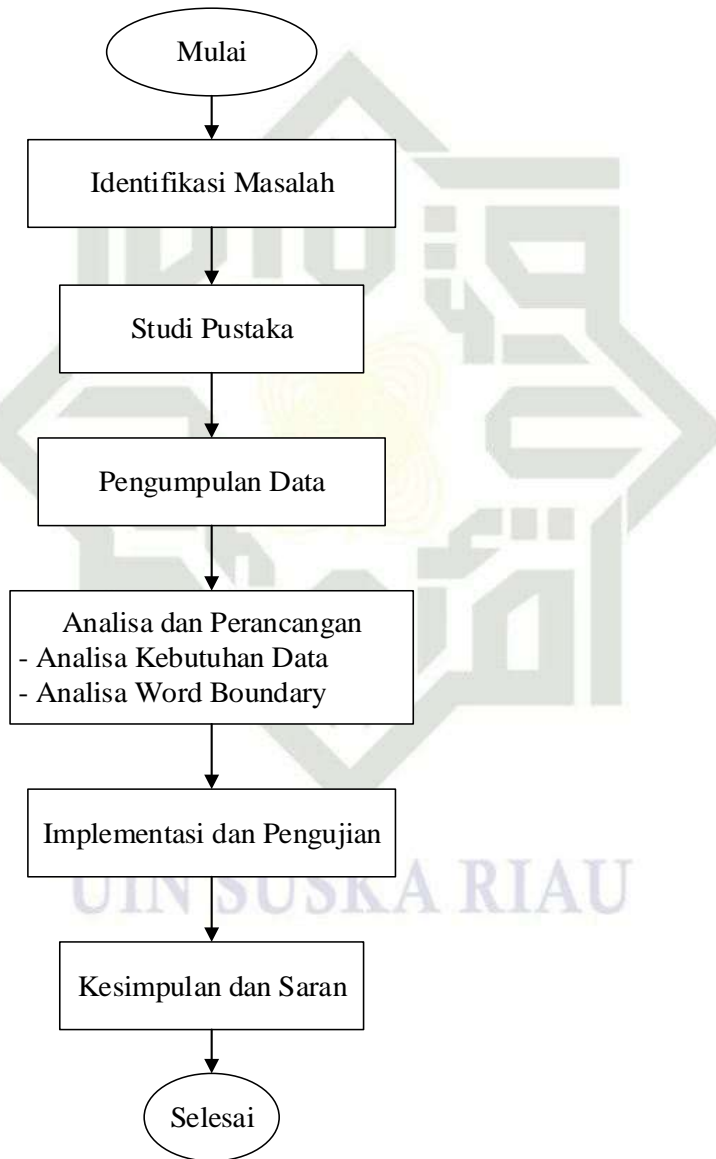
 <p>11.</p>	<p>Imam Habibi, Fahaldi</p>	<p>The Balinese Unicode Text Processing</p>	<p>Sistem ini telah berhasil dilakukan pengolahan teks Bali Unicode meliputi transliterasi, mencari, menyortir, dan kata fungsi batas. Selain itu, perpustakaan umum dan multiplatform telah berhasil dibangun. perpustakaan ini berjalan dengan baik pada kedua komputer dan perangkat genggam selama JVM (java mesin virtual) diinstal pada mereka</p>
<p>UIN SUSKA RIAU</p>	<p>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.</p>	<p>© Hak cipta milik UIN Suska Riau</p> <p>State Islamic University of Sultan Syarif Kasim Riau</p>	<p>Integrating Word Boundary Identification with Sentence Understanding</p> <p>Ambiguitas batas kata terdeteksi, khususnya yang salah, ketika mereka memblokir primer tugas menemukan hubungan antar berbagai konstituen kalimat, yang pada dasarnya adalah esensi dari proses pemahaman. Dalam pendekatan ini, informasi statistik juga dimasukkan, menyediakan sistem landasan awal yang cepat dan cukup andal untuk dibawa tugas utama membangun hubungan.</p>

UIN SUSKA RIAU



BAB III METODOLOGI PENELITIAN

Metodologi penelitian merupakan tahapan yang dilakukan untuk mendapatkan rancangan serta tujuan yang jelas tentang sistem yang akan dibangun. Adapun tahapan-tahapan yang dilakukan penulis dalam metode penelitian ini dapat dilihat pada gambar 3.1 berikut ini:



Gambar 3. 1 Flowchart Metodologi Penelitian

1. Di rangkai menjadi bagian-bagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. rangkaiannya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. rangkaiannya untuk mengikis kepentingan yang wajar UIN Suska Riau.
2. Di rangkai menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Berdasarkan *flowchart* pada Gambar 3.1 Metodologi Penelitian dalam

pengejaan tugas akhir meliputi beberapa tahapan yaitu :

3.1. Identifikasi Masalah

Identifikasi masalah merupakan tahapan awal dalam sebuah penelitian untuk menemukan topik permasalahan apa yang akan diangkat dalam penelitian ini, tujuan penelitian, dan menentukan batasan-batasan yang dibuat untuk membatasi penelitian ini agar tidak jauh menyimpang dari rumusan masalah dalam penelitian.

3.2. Studi Pustaka

Pada tahap studi pustaka dilakukan pencarian referensi atau informasi yang berhubungan dengan Identifikasi *Word-Boundary* pada *Hashtag Tweet* Bahasa Indonesia dan sistem yang akan dibangun.

3.3. Pengumpulan Data

Pada tahap pengumpulan data dilakukan pengumpulan data terkait yang dibutuhkan dalam penelitian. Ada beberapa jenis data yang dibutuhkan dalam penelitian ini, yaitu:

Data Hashtag

Data *hashtag* digunakan sebagai sumber data pengujian dari algoritma *Word-boundary*. Data *hashtag* diambil dari media sosial *Twitter* yang menggunakan *hashtag* sebanyak 500 *hashtag* yang menggunakan Bahasa Indonesia.

Kamus Bahasa Indonesia

Kamus digunakan sebagai sumber data untuk validasi hasil pengujian algoritma *Word-boundary* yang digunakan dalam penelitian. Pada tahapan ini hasil pemisahan kata dari data *hashtag* menggunakan algoritma *Word-boundary* yang di dapat akan di cek di kamus Bahasa Indonesia yang telah terdapat dalam *database*.

3.4. Analisa dan Perancangan

Setelah berbagai data serta informasi didapatkan dalam tahap pengumpulan data, maka tahap selanjutnya yaitu menganalisa permasalahan. Pada tahapan ini merupakan tahapan dimana peneliti mendalami serta menganalisa permasalahan yang terjadi.

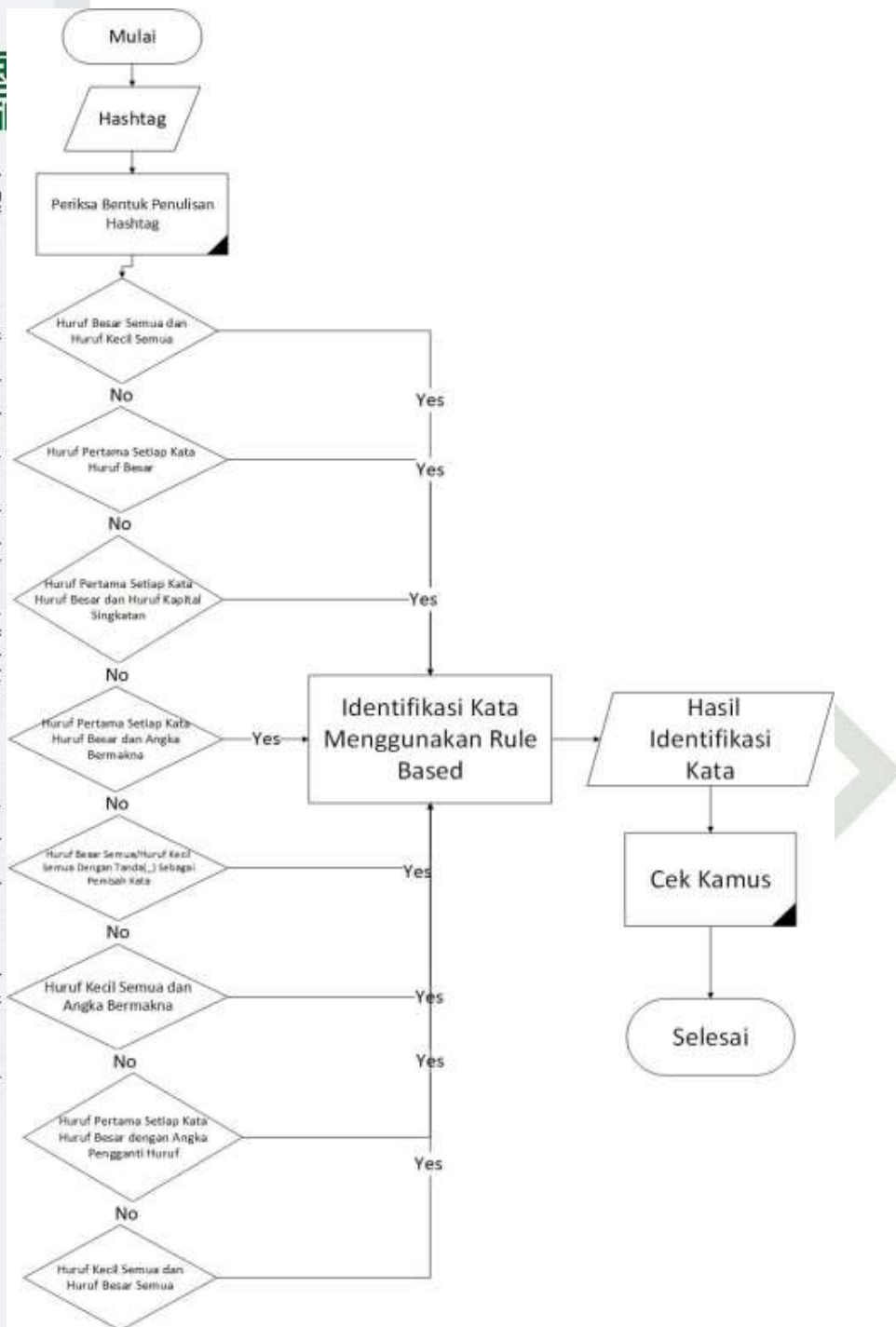
3.4.1. Analisa Kebutuhan Data

Analisa kebutuhan data merupakan tahapan menganalisa data *tweet* yang dikumpulkan dan kemudian akan digunakan pada penelitian ini. Analisa kebutuhan data membahas mengenai *tweet* dari media sosial *Twitter* yang memiliki tagar *hashtag* yang kemudian diproses. Namun data yang diproses merupakan teks yang terdapat pada tagar/*hashtag* saja. Dan ditulis menggunakan Bahasa Indonesia.

3.4.2. Analisa Word Boundary

Tahap selanjutnya yaitu analisa *word boundary* yang akan dilakukan pemeriksaan kesalahan ejaan dan kesalahan peletakan spasi pada teks. Contohnya teks pada *hashtag*. Apabila masih masih terdapat teks yang menyatu maka akan dilakukan *word boundary* menggunakan metode *rule based* untuk mendapatkan ejaan kata dan peletakan spasi yang benar.





Gambar 3. 2 Flowchart Analisa Identifikasi Word-Boundary

Berdasarkan *flowchart* diatas, maka dapat dilihat analisa proses *word-boundary* yaitu:

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 mengemukakan 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.
 2. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan satu masalah.
 3. Dilarang mengutip dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Setiap teks *hashtag* yang diiputkan akan diberi batas kata menggunakan aturan tata bahasa pada tahapan *Rule Based*. Setelah pemotongan kata berhasil maka akan mendapatkan kemungkinan solusi.

2. Setelah didapatkan kandidat solusi, dilakukan pemeriksaan kata pada kamus untuk memastikan kata-kata tersebut terdapat dalam kamus Bahasa Indonesia.

3. Setelah melakukan pemeriksaan pada kamus maka didapatkan hasil validasi pembenaran kata berdasarkan pada kamus Bahasa Indonesia.

3.5. Implementasi dan Pengujian

Setelah melakukan analisa dan perancangan, maka selanjutnya dilakukan implementasi dan pengujian untuk membangun sistem berdasarkan analisa permasalahan yang telah dibahas sebelumnya.

Perangkat keras yang digunakan dalam pembuatan serta penerapan aplikasi ini adalah sebagai berikut:

Processor : *Intel Core i5*

Memory : *RAM 8 GB*

Hardisk : *500 GB*

Tahapan pengujian yang dilakukan:

1. Pengujian *black box*

Pengujian yang dilakukan menggunakan metode pengujian *black box*. Metode pengujian *black box* memfokuskan pada kebutuhan fungsional sistem, metode ini digunakan untuk mengetahui apakah sistem telah berjalan sebagaimana mestinya.

2. Pengujian akurasi

Pengujian akurasi yaitu pengujian yang dilakukan untuk melakukan pengamatan hasil yang didapat pada algoritma di aplikasi yang telah dibangun.



BAB IV ANALISA DAN PERANCANGAN

4.1. Analisa dan Perancangan

Terdapat beberapa komponen analisa yang akan dijelaskan yaitu mengenai analisa permasalahan, pengumpulan data, proses algoritma Identifikasi *Word Boundary* pada teks Bahasa Indonesia dengan menggunakan tokenisasi metode *Rule Based* pada sub bab berikut ini:

4.1.1. Analisa Permasalahan

Permasalahan dalam penelitian ini bagaimana mendeteksi bentuk kata pada penulisan *hashtag* yang terdapat di *social media Twitter* berbahasa Indonesia yang sesuai dengan *rule based*. Mesin tidak dapat memproses kata yang salah seperti kesalahan pengetikan yang disengaja atau tidak disengaja, penggunaan tanda baca yang tidak sesuai dengan tata bahasa, kata yang disingkat-singkat tidak sesuai dengan aturan dan kombinasi huruf besar dan huruf kecil yang tidak beraturan.

Sebagai contoh yaitu pada teks yang memiliki *hashtag*, seperti ‘Selamat hari Selasa #harisantrinasional’, teks yang memiliki *hashtag* tersebut seharusnya dapat dipisah menjadi ‘hari santri nasional’, kesalahan seperti ini merupakan kesalahan penulisan yang disengaja. Kata yang menyatu di dalam *hashtag* tersebut tidak bisa dipisah menggunakan tokenisasi. Hal ini menyebabkan kesalahan dalam ejaan sehingga menyebabkan arti yang berbeda. Kesalahan tersebut juga dapat membuat mesin sulit mendeteksi kata dari sebuah teks. Diperlukan adanya proses pemisahan kata ulang menggunakan *word-boundary* pada kalimat tersebut untuk memisahkan teks yang tidak memiliki pemisah kata untuk mendapatkan makna sebenarnya.

4.1.2. Analisa Pengumpulan Data


Analisa pengumpulan data terdiri dari pengumpulan data *hashtag* dan pengumpulan data kamus yang akan dijelaskan sebagai berikut:

4.1.2.1 Pengumpulan data *hashtag*

Rada penelitian ini data yang di ambil bersumber dari *social media Twitter* berupa *hashtag* yang menjadi *trending topic* setiap harinya yang dikumpulkan hingga mencapai 500 data *hashtag*. Proses pengumpulan kata dasar dilakukan secara manual yaitu diketik ulang dengan memasukkan salinan *hashtag* yang telah dikumpulkan ke dalam *Microsoft Excel*. Berikut adalah tabel data *Hashtag*(#) yang digunakan dalam penelitian ini.

Tabel 4.1 Tabel Data *Hashtag*

no	Hashtag	keterangan
1	#SetasaRebahan	huruf pertama setiap kata huruf besar
2	#KemajuanNyataDiPapua	huruf pertama setiap kata huruf besar
3	#IisDahliaNorak	huruf pertama setiap kata huruf besar
4	#TanteMirna	huruf pertama setiap kata huruf besar
5	#DemoDibayarTapiAmbyar	huruf pertama setiap kata huruf besar
6	#TolakPemimpinKafir	huruf pertama setiap kata huruf besar
7	#TolakPemimpinWanita	huruf pertama setiap kata huruf besar
8	#SAPAMANTAN	huruf besar semua
9	#BeritaTerkini	huruf pertama setiap kata huruf besar
10	#beritaviral	huruf kecil semua
11	#DirutGarudaBaru	huruf pertama setiap kata huruf besar
12	#SenyumAja	huruf pertama setiap kata huruf besar
13	#UntukPapuaMaju	huruf pertama setiap kata huruf besar
14	#BoikotProdukCinaSekarang	huruf pertama setiap kata huruf besar
15	#PapuaPastiMakmur	huruf pertama setiap kata huruf besar

16		#lebahganteng	huruf kecil semua
		#bandunggeger	huruf kecil semua
		#haripatahhatinasional	huruf kecil semua
		#rabuambyar	huruf kecil semua
		#LiburanKeJogjaItuIstimewa	huruf pertama setiap kata huruf besar
		#akugapapakok	huruf kecil semua
.....			
500		#UngkapKebenaranKasusIndah	huruf pertama setiap kata huruf besar

Data selengkapnya pada lampiran A

Setelah dilakukan pengumpulan data *hashtag* maka didapatkan beberapa bentuk dalam penulisan sebuah *hashtag* di media sosial *Twitter*. Dari beberapa bentuk penulisan *hashtag* tersebut maka data dikategorikan sesuai dengan bentuk penulisan nya. Berikut adalah tabel bentuk penulisan *hashtag*:

Tabel 4. 2 Jenis Penulisan *Hashtag*

no	Contoh	format penulisan
1	#SAPAMANTAN	huruf besar semua
2	#hukummatikoruptor	huruf kecil semua
3	#IndonesiaDiAmbangKebangkrutan	huruf pertama setiap kata huruf besar
4	#PapuaMengecamOPM	huruf pertama setiap kata huruf besar dan huruf kapital singkatan
5	#Salam30Hari	huruf pertama setiap kata huruf besar dan mengandung angka bermakna

6	#4niesRajaKambingHitam	angka pengganti huruf, huruf pertama setiap kata huruf besar
7	#VIRUS_CORONA	huruf besar semua dan tanda baca () sebagai pemisah kata
8	#Bani_sumpah_pemuda	huruf kecil semua dan tanda baca () sebagai pemisah kata
9	#Sur14hari	huruf kecil semua dan angka bermakna
10	#JowoMUNDUR	huruf kecil semua dan huruf besar semua

Tabel diatas merupakan tabel data *hashtag* yang sudah di masukan ke dalam *Ms. Excel* yang didapat dari beberapa komentar yang ada diaplikasi *twitter*. Data yang sudah di masukkan ke dalam *Ms. Excel* tersebut kemudian akan di *export* ke dalam *database* untuk digunakan dalam pengujian pada sistem informasi yang dibangun.

4.1.2.2 Pengumpulan data kamus Bahasa Indonesia

Data kamus Bahasa Indonesia yang dikumpulkan bersumber dari Kamus Besar Bahasa Indonesia (KBBI) dan (Handayani dkk.). Fungsi dari data kamus adalah sebagai sumber pemeriksaan kebenaran dari hasil identifikasi *Word-boundary* pada data *hashtag* sebelumnya. Proses pengumpulan data kamus dilakukan secara manual yaitu diketik ulang dengan memasukkan salinan kata yang terdapat di buku kamus kedalam aplikasi *Microsoft Excel (Ms.Excel)*. Data kamus berisi dari kata-kata dalam Bahasa Indonesia, dan juga terdapat beberapa nama tokoh penting, nama tempat atau daerah, dan nama perusahaan. Berikut adalah tabel data kamus yang digunakan dalam penelitian ini:

Tabel 4. 3 Tabel Data Kamus

No	kata
1	a
2	ab
3	aba
4	aba=aba
5	abad
6	abadi
7	abadiah
8	abah
9	abah-abah
10	abai
11	abaimana
12	abaka
13	abaktinal
14	abakus
15	abal-abal
16	aban

31237	zuriat

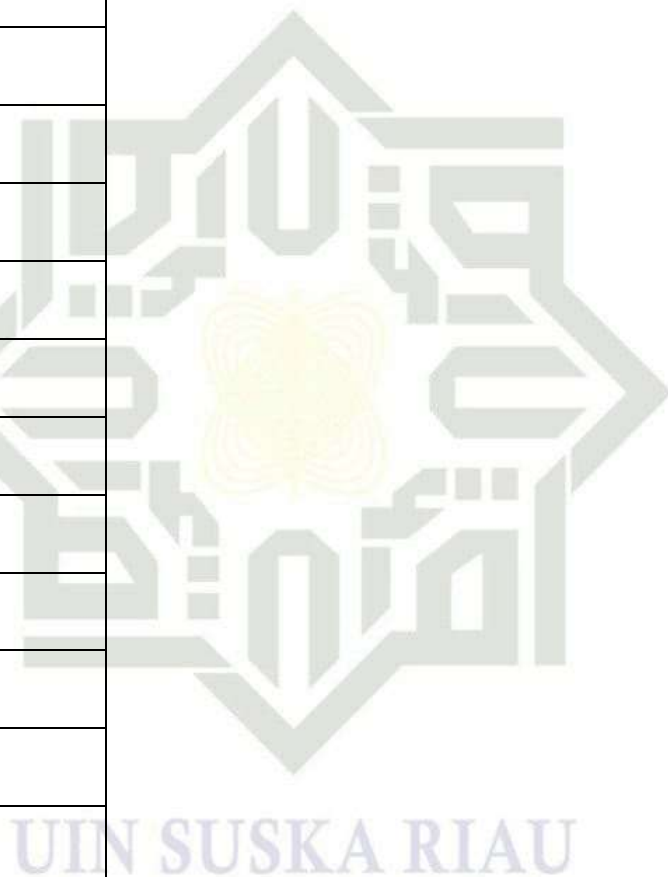
UIN Suska Riau



Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim 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, penerjemahan, 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.



Tabel diatas merupakan tabel data kamus yang sudah di masukan ke dalam Ms.excel yang didapat dari Kamus Besar Bahasa Indonesia. Total 31237 kata yang dimasukkan ke dalam Ms.excel kamus. Data yang sudah di masukan ke dalam Ms. Excel tersebut kemudian akan di export ke dalam database untuk digunakan dalam pengujian pada sistem informasi yang dibangun.

4.2. Word Boundary

Pada proses Word Boundary dilakukan pemeriksaan batas kata dalam teks yang ditunjukkan dengan adanya spasi antar kata. Maka pemeriksaan dilakukan terhadap data hashtag yang tidak memiliki batas kata sebagai berikut:

4.2.1. Pendekatan Rule Based

Setelah data hashtag dikumpulkan dan didapatkan beberapa bentuk penulisan hashtag. Dengan bentuk penulisan tanpa pembatas kata. Maka diberikan aturan penyelesaian agar data bisa diproses. Aturan dibuat berdasarkan dari bentuk penulisan hashtag yang telah didapat. Berdasarkan dari 10 jenis penulisan yang ada, terdapat beberapa jenis penulisan yang bisa diselesaikan dengan rule yang sama. Maka penulis memberikan aturan penyelesaian word-boundary pada hashtag menjadi 6 aturan, berdasarkan analisa dari 10 jenis penulisan hashtag yang telah didapatkan, sebagai berikut:

Menjadikan huruf kapital sebagai awalan kata

Setiap data yang memiliki bentuk penulisan “huruf pertama setiap kata huruf besar” akan dilakukan proses rule based dilihat dari bentuk data aturan yang dipakai adalah setiap huruf kapital merupakan awalan sebuah kata.

Contoh Penyelesaian

a. #SenyuminAja

#SenyuminAja

Senyumin Aja

Maka didapatkan rekomendasi kata = Senyumin, Aja



b. #UntukPapuaMaju

#UntukPapuaMaju

Untuk Papua Maju

Maka di dapatkan rekomendasi kata = Untuk, Papua, Maju

#IisDahliaNorak

#IisDahliaNorak

Iis Dahlia Norak

Maka di dapatkan rekomendasi kata = Iis, Dahlia, Norak

#BoikotProdukCinaSekarang

Boikot Produk Cina Sekarang

Maka di dapatkan rekomendasi kata = Boikot, Produk, Cina, Sekarang

#PapuaPastiMakmur

Papua Pasti Makmur

Maka di dapatkan kandidat kata = Papua, Pasti, Makmur

2. Memeriksa kumpulan huruf yang memiliki makna pada kamus sebagai kandidat sebuah kata

Setiap data yang memiliki bentuk penulisan “huruf besar semua” , “huruf kecil semua” dan “huruf kecil semua dan huruf besar semua” akan dilakukan proses *rule based* dilihat dari bentuk data aturan yang dipakai adalah dengan memeriksa huruf per huruf, dimana dicari kumpulan huruf yang memiliki makna pada kamus untuk dijadikan kandidat kata. Apabila kandidat kata ditemukan, maka dilanjutkan pencarian kata pada huruf selanjutnya.

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

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber.

a. Penelitian dan/atau karya tulis yang diterbitkan, penulisan karya ilmiah, penulisan laporan, penulisan skripsi, tesis, dan/atau disertasi pada jenjang pendidikan sarjana, pascasarjana, dan/atau penelitian tindakan.

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

c. Pengutipan untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan skripsi, tesis, dan/atau disertasi pada jenjang pendidikan sarjana, pascasarjana, dan/atau penelitian tindakan.

d. Pengutipan untuk kepentingan umum.

e. Pengutipan untuk kepentingan akademik.

f. Pengutipan untuk kepentingan jurnalistik.

g. Pengutipan untuk kepentingan hukum.

h. Pengutipan untuk kepentingan lain yang tidak bertentangan dengan ketentuan peraturan perundang-undangan.

i. Pengutipan untuk kepentingan lain yang tidak bertentangan dengan ketentuan peraturan perundang-undangan.

j. Pengutipan untuk kepentingan lain yang tidak bertentangan dengan ketentuan peraturan perundang-undangan.

k. Pengutipan untuk kepentingan lain yang tidak bertentangan dengan ketentuan peraturan perundang-undangan.

l. Pengutipan untuk kepentingan lain yang tidak bertentangan dengan ketentuan peraturan perundang-undangan.

m. Pengutipan untuk kepentingan lain yang tidak bertentangan dengan ketentuan peraturan perundang-undangan.

n. Pengutipan untuk kepentingan lain yang tidak bertentangan dengan ketentuan peraturan perundang-undangan.

o. Pengutipan untuk kepentingan lain yang tidak bertentangan dengan ketentuan peraturan perundang-undangan.

p. Pengutipan untuk kepentingan lain yang tidak bertentangan dengan ketentuan peraturan perundang-undangan.

q. Pengutipan untuk kepentingan lain yang tidak bertentangan dengan ketentuan peraturan perundang-undangan.

Contoh penyelesaian

UIN SUSKA RIAU



a. #lebahganteng

n=2

Hak Cipta Dilindungi Undang-Undang

Dilarang menyalin sebagian atau seluruh karya tulis ini tanpa ijin pencantuman dan menyebutkan sumber: a. Pengutipan harus mencantumkan kepenyertaan 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.

© Hachipriatmik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

le = tidak ada dikamus, sehingga n+1

eb = tidak ada dikamus, sehingga n+1

leb = tidak ada dikamus, sehingga n+1

lebah = ada dikamus, simpan sebagai kandidat

lanjut ke karakter berikutnya

n=2

Hak Cipta Dilindungi Undang-Undang

Dilarang menyalin sebagian atau seluruh karya tulis ini tanpa ijin pencantuman dan menyebutkan sumber: a. Pengutipan harus mencantumkan kepenyertaan 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.

© Hachipriatmik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

ga = tidak ada dikamus, sehingga n+1

gan = tidak ada dikamus, sehingga n+1

gant = tidak ada dikamus, sehingga n+1

ganteng = tidak ada dikamus, sehingga n+1

ganteng ada dikamus, simpan sebagai kandidat

maka didapatkan kandidat kata: “lebah”, “ganteng”.

b. #SAPAMANTAN

n=2

Hak Cipta Dilindungi Undang-Undang

Dilarang menyalin sebagian atau seluruh karya tulis ini tanpa ijin pencantuman dan menyebutkan sumber: a. Pengutipan harus mencantumkan kepenyertaan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah. b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

SA = tidak ada dikamus, sehingga n+1

SAP = tidak ada dikamus, sehingga n+1

SAPA = ada dikamus, simpan sebagai kandidat

lanjut ke karakter berikutnya



UIN SUSKA RIAU



n=2

MA = tidak ada dikamus, sehingga n+1

MAN = tidak ada dikamus, sehingga n+1

MANT = tidak ada dikamus, sehingga n+1

MANTA = tidak ada dikamus, sehingga n+1

MANTAN = ada dikamus, simpan sebagai kandidat

Maka didapatkan kandidat kata : “SAPA”, “MANTAN”.

#jokowiMUNDUR

n=2

o = tidak ada dikamus, sehingga n+1

ok = tidak ada dikamus, sehingga n+1

oko = tidak ada dikamus, sehingga n+1

okow = tidak ada dikamus, sehingga n+1

okowi = ada dikamus, simpan sebagai kandidat

anjut ke karakter berikutnya

n=2

MU = tidak ada dikamus, sehingga n+1

MUN = tidak ada dikamus, sehingga n+1

MUND = tidak ada dikamus, sehingga n+1

MUNDU = tidak ada dikamus, sehingga n+1

MUNDUR = ada dikamus, simpan sebagai kandidat.

Maka didapatkan kandidat kata : “jokowi”, “MUNDUR”.

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

Makalah Diindonesiakan oleh UIN Suska Riau
 Dilarang menyalin sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya.
 a. Pengutipan harus mencantumkan kepenyertaan, kepenyertaan pendahuluan, kepenyertaan penutup, kepenyertaan laporan, kepenyertaan kritik atau tinjauan, kepenyertaan hasrat.
 b. Pengutipan tidak mengizinkan kepenyertaan yang wajar UIN Suska Riau.





Hak Cipta milik UIN Suska Riau
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumbernya.
2. Dilarang menggunakan sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Staf Islamic University of Sultan Syarif Kasim Riau

3. Menjadikan beberapa huruf kapital berurutan sebagai singkatan

Setiap data yang memiliki bentuk penulisan “huruf pertama setiap kata huruf besar dan huruf capital singkatan” akan dilakukan proses *rule based* dilihat dari bentuk data, aturan yang dipakai adalah dengan memeriksa huruf per huruf, dimana setiap huruf kapital yang dilanjutkan dengan huruf biasa merupakan awalan sebuah kata namun apabila huruf kapital tersebut dilanjutkan dengan huruf kapital lagi maka dianggap sebagai penulisan singkatan.

Contoh penyelesaian:

#ILCSalahkanAnies
ILCSalahkanAnies
ILCSalahkanAnies
ILCSalahkanAnies

Maka didapatkan kandidat kata = “Salahkan”, “Anies”

ILC = kapital+kapital+kapital, dianggap kandidat singkatan

Apabila terdeteksi kata, maka huruf kapital yang dilanjutkan huruf biasa dianggap awalan kata, kandidat kata berhenti sampai terdeteksi huruf kapital selanjutnya.

Maka didapatkan penyelesaian singkatan+kata = “ILC”, “Salahkan”, “Anies”.

4. Menjadikan tanda baca sebagai pemisah kata

Setiap data yang memiliki bentuk penulisan “huruf besar semua dan tanda baca (.) sebagai pemisah” dan “huruf kecil semua dan tanda baca () sebagai pemisah” maka akan dilakukan proses *rule based* dilihat dari bentuk data aturan yang dipakai adalah setiap huruf yang ditulisa berurutan merupakan kata dengan tanda baca sebagai pemisah kata.

Contoh penyelesaian:

UIN SUSKA RIAU



a. #VIRUS_CORONA

tanda baca () sebagai pemisah kata = “VIRUS”, “CORONA”.

#harisumpah_pemuda

tanda baca () sebagai pemisah kata = “hari”, ”sumpah”, ”pemuda”.

5. Simpan penulisan angka sebagai angka mengandung makna

Setiap data yang memiliki bentuk penulisan “huruf pertama setiap kata huruf besar mengandung angka bermakna” dan “huruf kecil semua dan angka bermakna” maka akan dilakukan proses *rule based* dilihat dari bentuk data. Aturan yang dipakai adalah dengan memeriksa huruf per huruf, dimana setiap huruf kapital yang dilanjutkan dengan huruf biasa merupakan awalan sebuah kata, dan apabila terdapat penulisan angka maka disimpan sebagai angka bermakna, seperti tanggal ataupun hari besar. Namun apabila tidak terdapat huruf kapital maka diproses dengan memeriksa huruf per huruf, dimana dicari kumpulan huruf yang memiliki makna pada kamus untuk dijadikan kandidat kata. Apabila kandidat kata ditemukan, maka dilanjutkan pencarian kata pada huruf selanjutnya.

Contoh penyelesaian:

#BanjirJakarta2020

Banjir Jakarta 2020

Maka di dapatkan kandidat kata = “Banjir”, ”Jakarta”, ”2020”.

#libur14hari

N=2

li = tidak ada dikamus, sehingga n+1

lib = tidak ada dikamus, sehingga n+1

Hak cipta milik UIN Suska Riau
Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengizinkan penulisnya untuk membuat laporan penelitian, pengajaran, atau pelatihan. Penyalinan untuk kepentingan pribadi diperbolehkan.
Dilarang mengutip dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



libu = tidak ada dikamus, sehingga n+1

libur = ada dikamus, simpan sebagai kandidat

lanjut ke karakter berikutnya

n=2

14 = angka bermakna, simpan sebagai kandidat

lanjut ke karakter berikutnya

n=2

ha = tidak ada dikamus, sehingga n+1

har = tidak ada dikamus, sehingga n+1

hari = ada dikamus, simpan sebagai kandidat

Maka didapatkan kandidat kata : “libur”, “14”, “hari”.

Menjadikan angka sebagai pengganti huruf dalam kata

Setiap data yang memiliki bentuk penulisan “angka pengganti huruf dan

huruf pertama setiap kata huruf besar” akan dilakukan proses *rule based* dilihat

dari bentuk data aturan yang dipakai adalah setiap huruf kapital merupakan

awal sebuah kata dan penulisan angka sebagai pengganti huruf dalam kata.

#4niesRajaKambingHitam

4nies Raja Kambing Hitam

Maka didapatkan kandidat kata : “4nies”, “Raja”, “Kambing”, “Hitam”.

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

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 tesis atau tinjauan suatu masalah.

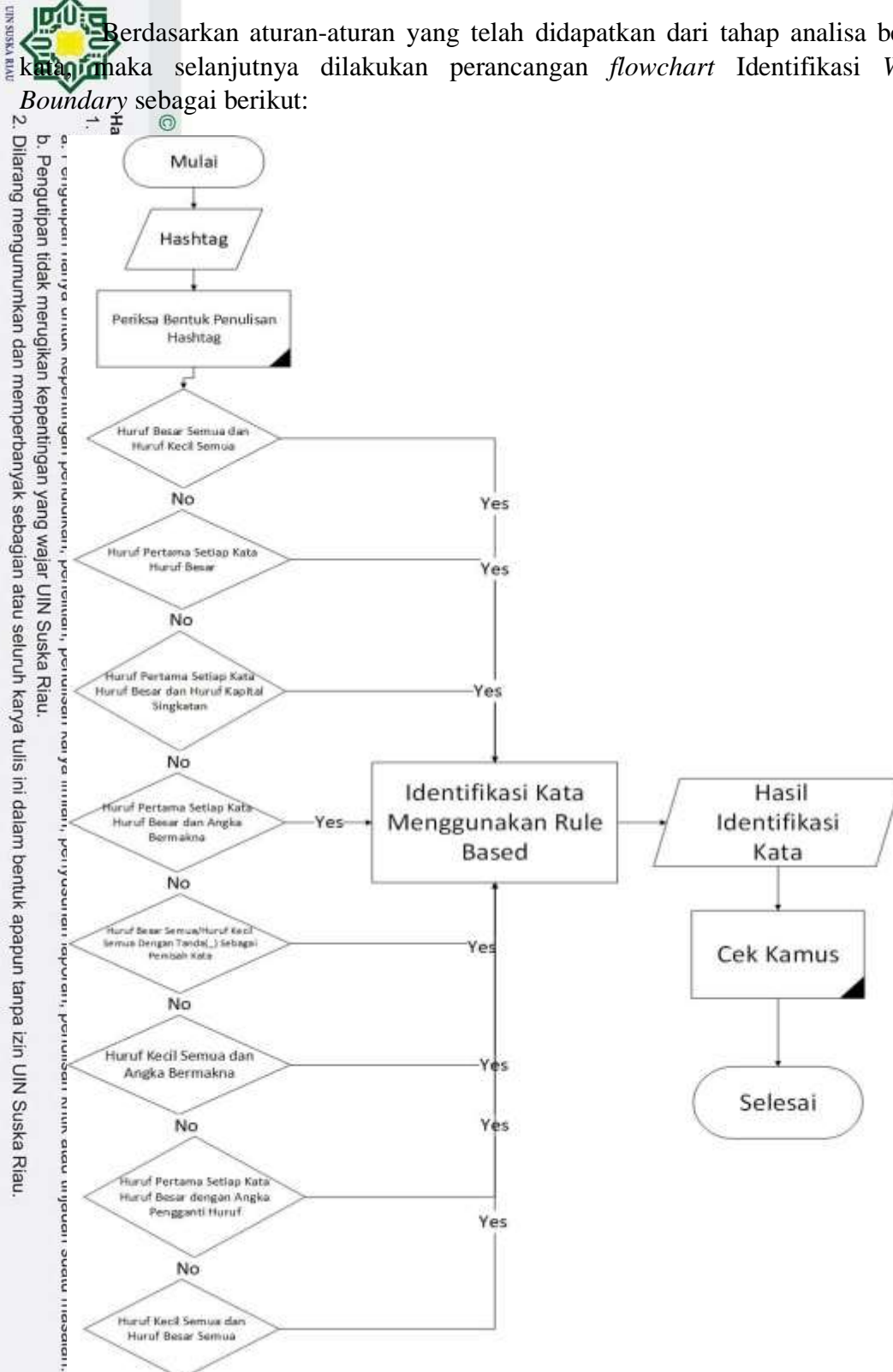
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

4.2.2. Analisa Flowchart Identifikasi Word-boundary

Berdasarkan aturan-aturan yang telah didapatkan dari tahap analisa bentuk kata, maka selanjutnya dilakukan perancangan *flowchart* Identifikasi *Word-Boundary* sebagai berikut:



Gambar 4. 1 Analisa *Flowchart* Identifikasi *Word-Boundary*

Penjelasan tahapan *flowchart*:

UIN SUSKA RIAU



1. Proses dimulai ketika memasukkan *hashtag*

2. *Hashtag* yang telah dimasukkan kemudian diperiksa bentuk penulisannya agar diproses sesuai aturan yang telah dibuat.

Jika *hashtag* memiliki bentuk penulisan “huruf besar semua/huruf kecil semua” maka dilakukan proses Identifikasi *Word-boundary* menggunakan *Rule-Based* yang cocok. Jika bukan, maka periksa ke bentuk selanjutnya.

Jika *hashtag* memiliki bentuk penulisan “huruf pertama setiap kata huruf besar” maka dilakukan proses Identifikasi *Word-boundary* menggunakan *Rule-Based* yang cocok. Jika bukan, maka periksa ke bentuk selanjutnya.

Jika *hashtag* memiliki bentuk penulisan “huruf pertama setiap kata huruf besar dan huruf kapital singkatan” maka dilakukan proses Identifikasi *Word-boundary* menggunakan *Rule-Based* yang cocok. Jika bukan, maka periksa ke bentuk selanjutnya.

Jika *hashtag* memiliki bentuk penulisan “huruf pertama setiap kata huruf besar dan angka bermakna” maka dilakukan proses Identifikasi *Word-boundary* menggunakan *Rule-Based* yang cocok. Jika bukan, maka periksa ke bentuk selanjutnya.

Jika *hashtag* memiliki bentuk penulisan “huruf besar semua/huruf kecil semua dengan tanda () sebagai pemisah kata” maka dilakukan proses Identifikasi *Word-boundary* menggunakan *Rule-Based* yang cocok. Jika bukan, maka periksa ke bentuk selanjutnya.

Jika *hashtag* memiliki bentuk penulisan “huruf kecil semua dan angka bermakna” maka dilakukan proses Identifikasi *Word-boundary* menggunakan *Rule-Based* yang cocok. Jika bukan, maka periksa ke bentuk selanjutnya.

Jika *hashtag* memiliki bentuk penulisan “huruf pertama setiap kata huruf besar dengan angka pengganti huruf” maka dilakukan proses Identifikasi

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

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.

UIN SUSKA RIAU
Hak cipta dilindungi undang-undang
Stre Isamic University of Sultan Saifuddin Kasir Riau



Word-boundary menggunakan *Rule-Based* yang cocok. Jika bukan, maka periksa ke bentuk selanjutnya.

10. Jika *hashtag* memiliki bentuk penulisan “huruf kecil semua dan huruf besar semua” maka dilakukan proses Identifikasi *Word-boundary* menggunakan *Rule-Based* yang cocok.

1. Lalu mendapatkan hasil identifikasi kata.

2. Lalu cek ketersediaan kata pada kamus. Untuk membuktikan kumpulan kata yang didapatkan merupakan identifikasi kata yang benar. Dan tahapan selesai.

4.2.3. Analisa Algoritma Identifikasi *Word-boundary*

Untuk mendapatkan penjelasan penyelesaian *word-boundary* maka dibutuhkan algoritma, yaitu langkah – langkah logis tertentu untuk menyelesaikan suatu masalah. Berikut merupakan algoritma untuk identifikasi *word-boundary* pada *hashtag tweet* bahasa Indonesia:

1. Mulai proses identifikasi dengan aturan *Rule Based*

Setelah mengumpulkan data teks *hashtag* yang memiliki bentuk penulisan terdapat huruf biasa dan huruf kapital sebagai awalan kata. Dan mengumpulkan data kamus Bahasa Indonesia. Maka dilanjutkan dengan proses identifikasi *word-boundary* dengan memberikan batas kata pada teks hashtag tersebut sesuai aturan tata bahasa dengan menjadikan setiap huruf kapital merupakan awalan sebuah kata.

2. Mendapatkan kemungkinan solusi kata

Setelah melakukan proses identifikasi *word-boundary*, maka akan mendapatkan kemungkinan solusi kata pada kolom hasil identifikasi *word-boundary*.

Hak cipta dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumbernya.
a. Pengutipan hanya untuk kepentingan pendidikan.
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.

Tabel 4. 4 Tabel Hasil Identifikasi Kata

No	Hastag Awal	Hasil
1	#hukummatikoruptor	hukum mati koruptor
2	#IndonesiaDiAmbangKebangkrutan	Indonesia Di Ambang Kebangkrutan
3	#RakyatBukanSapiPerah	Rakyat Bukan Sapi Perah
4	#ILCSalahkanAnies	ILC Salahkan Anies
5	#PapuaMengecamOPM	Papua Mengecam OPM
6	#BanjirJakarta2020	Banjir Jakarta 2020
7	#Salam30Hari	Salam 30 Hari
8	#4niesDatangIndonesiaSenang	4nies Datang Indonesia Senang
9	#4niesRajaKambingHitam	4nies Raja Kambing Hitam
10	#VIRUS_CORONA	VIRUS CORONA

2. Melakukan cek ketersediaan kata pada kamus

Setelah mendapatkan hasil dari proses identifikasi *word-boundary*, maka didapatkan beberapa kata hasil dari proses pemisahan bentuk teks menjadi beberapa bentuk kata. Lalu kata-kata tersebut diperiksa di kamus untuk melihat ketersediaan kata yang benar di kamus. Apabila berhasil ditemukan di kamus maka diberi keterangan pada kolom kamus. Berikut tabel periksa ketesediaan kata:

Tabel 4. 5 Tabel Periksa Ketersediaan Kata

No	Hastag Awal	Hasil	Kata
1	#hukummatikoruptor	hukum mati koruptor	hukum ada, mati ada, koruptor ada,
2	#IndonesiaDiAmbangKebangkrutan	Indonesia Di Ambang Kebangkrutan	indonesia ada, di ada, ambang ada, kebangkrutan ada,
3	#RakyatBukanSapiPerah	Rakyat Bukan Sapi Perah	rakyat ada, bukan ada, sapi ada, perah ada,
4	#ILCSalahkanAnies	ILC Salahkan Anies	ilc tidak ada, salahkan tidak ada, anies ada,
5	#PapuaMengecamOPM	Papua Mengecam OPM	papua ada, mengecam tidak ada, opm tidak ada,
6	#BanjirJakarta2020	Banjir Jakarta 2020	banjir ada, jakarta ada, 2020 tidak ada,
7	#Salam30Hari	Salam 30 Hari	salam ada, 30 tidak ada, hari ada,
8	#4niesDatangIndonesiaSenang	4nies Datang Indonesia Senang	4nies tidak ada, datang ada, indonesia ada, senang

			ada,
9	#4niesRajaKambingHitam	4nies Raja Kambing Hitam	4nies tidak ada, raja ada, kambing ada, hitam ada,
10	#VIRUS_CORONA	VIRUS CORONA	virus ada, corona ada,

4.2.4. Analisa Pseudocode Identifikasi Word-boundary

Analisa identifikasi *word-boundary* menggunakan *rule based* ditulis dalam bentuk *pseudocode* untuk dimasukkan ke dalam sistem. Pada perancangan *pseudocode* dimulai dari memasukkan data hashtag dengan *actionImport*. Bentuk *pseudocode* sebagai berikut:

```

BEGIN
function action import{
panggil model import
model import add rule fileImport->required
  if request post
    if model import->fileImport && modelImport->validate()
      set value baseRow =4
      while cek kolom B yang tidak kosong
        panggil model kamus
        masukkan nilai kolom B ke dalam model->kata
        masukkan nilai kolom C ke dalam model->format penulisan
        masukkan ke dalam database
        baseRow = baseRow+1
      endwhile
      show flashMessage(sukses)
    endif
  else
    show flashMessage(gagal)
  endif
endif
}
END

```

Gambar 4. 2 Pseudocode Import Hashtag

Setelah melakukan *import hashtag*, maka langkah selanjutnya yaitu melakukan identifikasi *import kamus* untuk melihat ketersediaan kata dari hasil identifikasi kata. berikut adalah bentuk *pseudocode*:

```

action import kamus

BEGIN
function action import{
panggil model import
model import add rule fileImport->required
  if request post
    if model import->fileImport && modelImport->validate()
      set value baseRow =4
      while cek kolom B yang tidak kosong
        panggil model kamus
        masukkan nilai kolom B ke dalam model->kamus
        masukkan ke dalam database
        baseRow = baseRow+1
      endwhile
      show flashMessage(sukses)
    endif
  else
    show flashMessage(gagal)
  endif
endif
}
END

```

Gambar 4. 3 Pseudocode Import Kamus

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 mengizinkan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengutip dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Setelah melakukan *import* kamus, maka langkah selanjutnya yaitu melakukan identifikasi *word-boundary* dengan menghapus bentuk penulisan selain huruf dan angka, seperti simbol titik, koma, bintang, pagar dll. Lalu juga diberi pembatas kata untuk mendapatkan hasil kata teridentifikasi sesuai dengan bentuk penulisan data *hashtag*. berikut adalah bentuk *pseudocode*:

```

BEGIN
    if kategori hashtag == "huruf kecil semua"
        hilangkan tanda pagar
        pisahkan semua huruf menjadi array

        for pisahkan semua huruf menjadi array => value
            cari di database setiap value
        endfor
    endif
END

```

Gambar 4. 4 Pseudocode Bentuk Penulisan Huruf Kecil Semua

Pseudocode diatas merupakan action untuk bentuk penulisan *hashtag* “huruf kecil semua” dilakukan sesuai aturan yang telah dibuat.

```

BEGIN
    if kategori hashtag == "huruf besar semua"
        jadikan setiap huruf menjadi huruf kecil
        hilangkan tanda pagar
        pisahkan semua huruf menjadi array

        for pisahkan semua huruf menjadi array => value
            cari di database setiap value
        endfor
    endif
END

```

Gambar 4. 5 Pseudocode Bentuk Penulisan Huruf Besar Semua

Pseudocode diatas merupakan action untuk bentuk penulisan *hashtag* “huruf besar semua” dilakukan sesuai aturan yang telah dibuat.

```

BEGIN
    if kategori hashtag == "huruf pertama setiap kata huruf besar"
        hapus tanda pagar
        pisahkan dengan aturan regex ('#[A-Z][a-z]*#')
        kemudian hasil dijadikan string
    endif
END

```

Gambar 4.6 Pseudocode Bentuk Penulisan Huruf Pertama Setiap Kata Huruf Besar



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

b. Pengutipan tidak mengizinkan kepada orang lain yang diperbolehkan untuk memperbanyak atau menyalin, mendistribusikan, mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Pseudocode diatas merupakan action untuk bentuk penulisan *hashtag* “huruf pertama setiap kata huruf besar” dilakukan sesuai aturan yang telah dibuat.

```

BEGIN
    .if kategori hashtag == "huruf pertama setiap kata huruf besar dan huruf kapital singkatan"
        hapus tanda pagar
        pisahkan dengan aturan regex ('#[A-Z][a-z]*#')
        kemudian hasil dijadikan string
        jadikan array dengan parameter spasi
        for pisahkan semua huruf menjadi array -> value
            if panjang string == 1 then
                pisahkan huruf
            else
                satukan huruf
            endif
        endfor
    endif
END
    
```

Gambar 4.7 Pseudocode Bentuk Penulisan Huruf Pertama Setiap Kata Huruf Besar dan Kapital Singkatan

Pseudocode diatas merupakan action untuk bentuk penulisan *hashtag* “huruf pertama setiap kata huruf besar dan kapital singkatan” dilakukan sesuai aturan yang telah dibuat.

```

BEGIN
    if kategori hashtag == "huruf pertama setiap kata huruf besar dan mengandung angka"
        hapus tanda pagar
        pisahkan dengan aturan regex ('#[A-Z][a-z]*#')
        kemudian hasil dijadikan string
    endif
END
    
```

Gambar 4.8 Pseudocode Bentuk Penulisan Huruf Pertama Huruf Besar dan Mengandung Angka

Pseudocode diatas merupakan action untuk bentuk penulisan *hashtag* “huruf pertama setiap kata huruf besar dan mengandung angka” dilakukan sesuai aturan yang telah dibuat.



1. Dilarang mengutip sebagian atau seluruh karya atau bagian dari karya tersebut, baik dengan cara apa pun, untuk tujuan komersial atau untuk kepentingan pribadi atau organisasi tanpa izin dari pihak yang berhak. 2. Dilarang mengutip sebagian atau seluruh karya atau bagian dari karya tersebut, baik dengan cara apa pun, untuk tujuan komersial atau untuk kepentingan pribadi atau organisasi tanpa izin UIN Suska Riau.

```
BEGIN
    if kategori hashtag == "huruf kecil semua dan tanda baca (_) sebagai pemisah kata "
        hilangkan tanda pagar
        pisahkan dengan aturan regex ('#[A-Z][a-z]*#')
        satukan menjadi string
        pisahkan tanda "_" sehingga menjadi array
        satukan array tersebut menjadi string
    endif
END
```

Gambar 4. 9 Pseudocode Bentuk Penulisan Huruf Kecil Semua dan Tanda Baca () Sebagai Pemisah Kata

Pseudocode diatas merupakan action untuk bentuk penulisan *hashtag* “huruf kecil semua dan tanda baca () sebagai pemisah kata” dilakukan sesuai aturan yang telah dibuat.

```
BEGIN
    if kategori hashtag == "huruf besar semua dan tanda baca(_) sebagai pemisah kata"
        hilangkan tanda pagar
        pisahkan dengan aturan regex ('#[A-Z][a-z]*#')
        satukan menjadi string
        pisahkan tanda "_" sehingga menjadi array
        satukan array tersebut menjadi string
    endif
END
```

Gambar 4. 10 Pseudocode Bentuk Penulisan Huruf Besar Semua dan Tanda Baca () Sebagai Pemisah Kata

Pseudocode diatas merupakan action untuk bentuk penulisan *hashtag* “huruf besar semua dan tanda baca () sebagai pemisah kata” dilakukan sesuai aturan yang telah dibuat.


```

BEGIN
  if kategori hashtag == "angka pengganti huruf, huruf pertama setiap kata huruf besar"
    hapus tanda pagar
    pisahkan dengan aturan regex ('#[A-Z][a-z]*#')
    kemudian hasil dijadikan string

    for pisahkan semua huruf menjadi array => value
      if panjang string == 1 then
        satukan huruf
      else
        pisahkan huruf
      endif
    endfor
  endif
END

```

Gambar 4.11 Pseudocode Bentuk Penulisan Angka Pengganti Huruf, Huruf Pertama Setiap Kata huruf Besar

Pseudocode diatas merupakan action untuk bentuk penulisan *hashtag* "angka pengganti huruf, huruf pertama setiap kata huruf besar" dilakukan sesuai aturan yang telah dibuat.

4.3. Perancangan Sistem

Perancangan Identifikasi *Word-Boundary* pada *hashtag tweet* Bahasa Indonesia dirancang berdasarkan tahapan analisa yang telah dilakukan sebelumnya. Bagian perancangan berisikan pembahasan mengenai perancangan basis data dan perancangan antarmuka.

4.3.2. Perancangan Basis Data

Basis Data adalah tempat penyimpanan data dalam bentuk digital. Basis Data menjadi salah satu kunci agar sistem yang dibangun berjalan sebagai sesuai keinginan. Perancangan Basis Data pada studi kasus ini adalah:

a. *Data Base User*

Berikut adalah rancangan *database* dari Tabel *User*

Tabel 4.6 Tabel User Pada Database

Name	Type	Length	Keterangan
Id	int	11	Primeri Key
username	Varchar	255	Unique
auth_key	Varchar	32	-

password_hash	Varchar	255	-
email	Varchar	255	Unique
status	smallint	6	
created_at	int	11	-
updated_at	int	11	-
verification_token	Varchar	255	-

Data Base Hashtag

Berikut adalah rancangan database dari tabel *Hashtag*:

Tabel 4. 7 Tabel *Hashtag* Pada *Database*

Name	Type	Length	Keterangan
id	int	11	Primeri Key
hashtag	Varchar	255	-

Data Base Hasil

Berikut adalah rancangan *database* dari Tabel awal *Hashtag* dan Hasil

Hashtag:

Tabel 4. 8 Tabel Hasil pada *Database*

Name	Type	Length	Keterangan
id	int	11	Primeri Key
awal_hashtag	Varchar	255	-
hasil_hashtag	Varchar	255	-

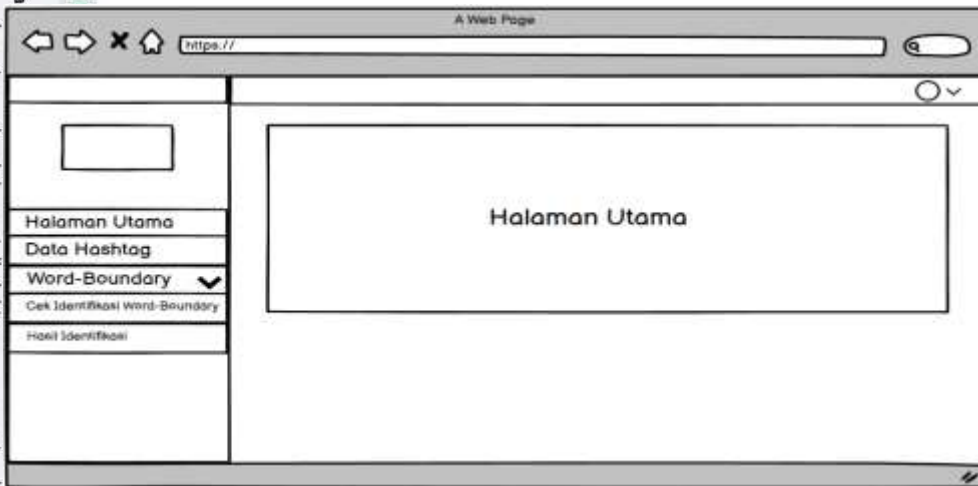
4.3.3. Perancangan *Interface*

Interface (antarmuka) adalah tampilan yang memudahkan *user* berinteraksi dengan sistem. Pada penelitian ini, perancangan *interface* dirancang menggunakan

aplikasi *Peniel*. Perancangan sistem ini terdapat beberapa *interface* yaitu:

1. Halaman Utama

Berikut adalah rancangan *interface* untuk halaman utama. Halaman ini dapat diakses oleh admin jika admin telah berhasil *login*.



Gambar 4. 12 *Interface* Halaman Utama

Halaman Data *Hashtag*

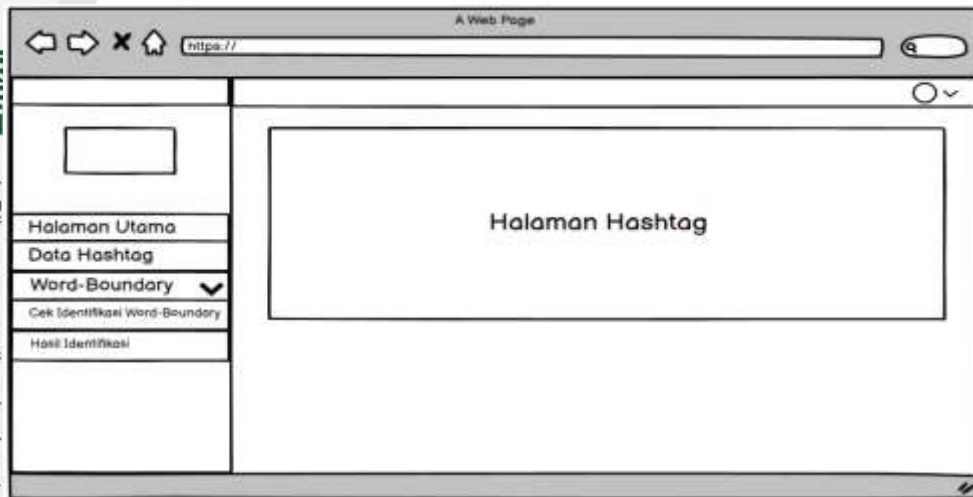
Berikut adalah rancangan *interface* untuk untuk halaman data *hashtag*.

Halaman ini dapat diakses oleh admin jika admin telah berhasil *login*.

Halaman data *hashtag* berisi tabel yang isi dari tabel tersebut data *hashtag* yang tersimpan pada *database*.



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.



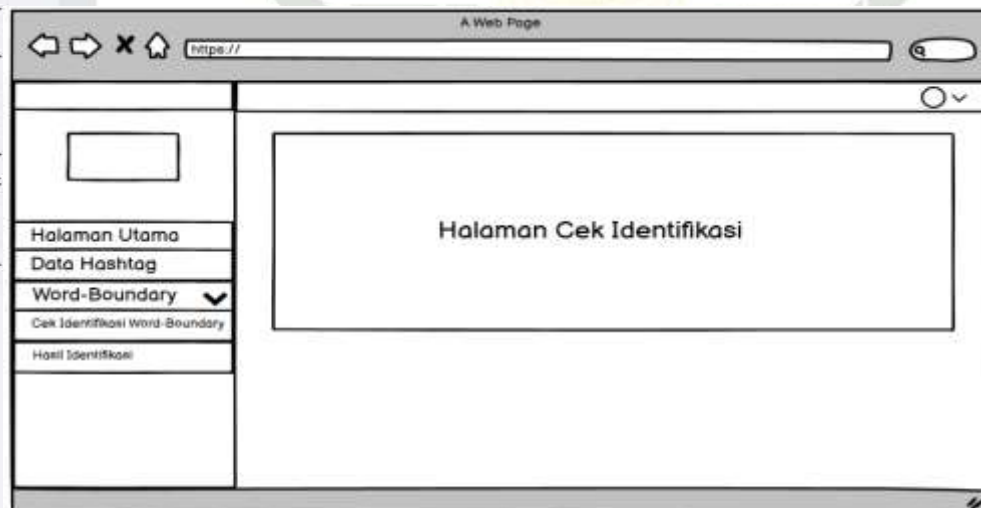
Gambar 4. 13 Interface Halaman Data Hashtag

Halaman Tes Identifikasi

Berikut adalah rancangan *interface* untuk untuk halaman test identifikasi.

Halaman ini dapat diakses oleh admin jika admin telah berhasil *login*.

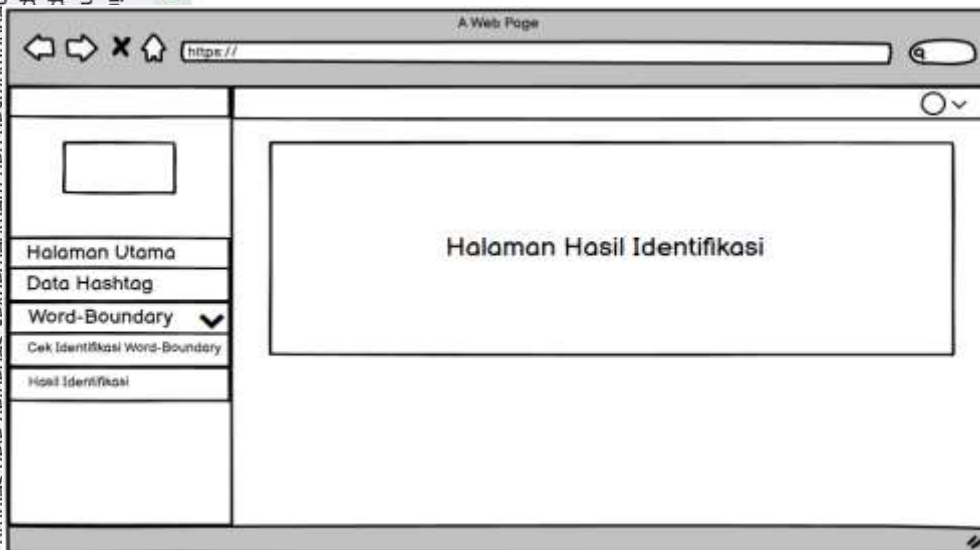
Halaman data *hashtag* berisi form untuk mengidentifikasi sebuah *hashtag* yang akan tersimpan pada halaman hasil identifikasi.



Gambar 4. 14 Interface Tes Identifikasi

d. Halaman Hasil Identifikasi

Berikut adalah rancangan *interface* untuk untuk halaman hasil identifikasi. Halaman ini dapat diakses oleh admin jika admin telah berhasil *login*. Halaman ini berisi hasil dari inputan pada halaman test identifikasi *hashtag*. Pada halaman ini berisi tabel dari hasil inputan pada halaman test identifikasi yang akan tersimpan pada *database*.



Gambar 4. 15 *Interface* Hasil Identifikasi

UIN SUSKA RIAU
2. Ditaring Mengumpulkan dari internet yang sekarang atau dari lain nya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

UIN SUSKA RIAU
2. Ditaring Mengumpulkan dari internet yang sekarang atau dari lain nya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

UIN SUSKA RIAU
2. Ditaring Mengumpulkan dari internet yang sekarang atau dari lain nya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

UIN SUSKA RIAU
2. Ditaring Mengumpulkan dari internet yang sekarang atau dari lain nya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



BAB VI

KESIMPULAN DAN SARAN

6.1. Kesimpulan

Berdasarkan hasil penelitian tentang algoritma Identifikasi *Word-Boundary* pada *Hashtag* Tweet Bahasa Indonesia dapat disimpulkan:

1. Hasil identifikasi kata terhadap *hashtag* berhasil dijalankan sesuai perancangan dalam hampir seluruh keluaran dari sistem sesuai dengan harapan yang diinginkan. Dimana pengujian akurasi terhadap 500 *hashtag* didapatkan akurasi sebesar 98,2% dengan 10 jenis kombinasi algoritma *rule based*. Berdasarkan dari 10 jenis penulisan *hashtag*, hanya 1 jenis penulisan yang masih terdapat hasil Identifikasi *Word-Boundary* yang salah, dengan jumlah akurasi per jenis sebesar 92%, sementara 9 jenis lainnya masing-masing memiliki jumlah akurasi sebesar 100%. Secara garis besar dengan nilai akurasi yang didapat, menyatakan bahwa algoritma dapat berjalan dengan baik.
2. Untuk data *hashtag* yang menggunakan nama orang, nama tempat, nama daerah dan nama perusahaan lebih baik menambahkan data tersebut ke dalam kamus agar kata dapat teridentifikasi oleh kamus.

6.2. Saran

Berdasarkan hasil kesimpulan dalam penelitian, terdapat saran yang dapat digunakan untuk penelitian selanjutnya yang lebih baik sebagai berikut:

1. Pada penelitian selanjutnya diharapkan peneliti dapat menggunakan pencarian berulang agar didapatkan hasil kata yang lebih akurat.
2. Semakin banyak aturan yang dirancang, maka akan semakin banyak tingkat keberhasilan proses identifikasi *Word-Boundary*.

DAFTAR PUSTAKA

- Asri, L. W., Rachmat C., A., & Lukito, Y. (2017). *Implementasi Algoritma Naïve Bayes Menggunakan Isear Untuk Klasifikasi Emosi Lirik Lagu Berbahasa Inggris*. *Jurnal Informatika*, 14(1), 16–21. <https://doi.org/10.9744/informatika.14.1.16-21>
- Falla, A. . (2013). *Penggunaan Hashtag (#) Di Media Sosial Untuk Advokasi Kesadaran Membayar Pajak Using Hashtag (#) In Social Media For Taxpaying Awareness*. (September), 0–10.
- Habibi, I. (2009). *The Balinese Unicode Text Processing*. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 1(1). <https://doi.org/10.22146/ijccs.19>
- Huang, C.-R., Simon, P., Hsieh, S., & Pr, L. (2007). *Rethinking Chinese Word Segmentation : Tokenization , Character Classification , or Wordbreak Identification*. *Computational Linguistics*, (June), 69–72. Retrieved from <http://www.aclweb.org/anthology/P07-2018>
- Juditha, C. (2015). *Fenomena Trending Topic Di Twitter : Analisis Wacana Twit # Savehajilulung Trending Topic Phenomenon On Twitter : Discourse Analysis Of Tweet # Savehajilulung*. 16(2), 138–154.
- Jurish B. (2013). *Word and Sentence Tokenization with Hidden Markov Models*. *Journal for Language Technology and Computational Linguistics*, 28(2), 61–88.
- Koehn, P., & Knight, K. (2003). *Empirical methods for compound splitting*. 187. <https://doi.org/10.3115/1067807.1067833>
- Lisangan, E. A. (2015). *Natural Language Processing Dalam Memperoleh Informasi Akademik Mahasiswa*. *Jurnal Tematika*, 1(March 2013), 1–9.
- Mahfira, T. N., Cholissodin, I., & Widodo, A. W. (2017). *Deteksi Kesalahan Ejaan dan Penentuan Rekomendasi Koreksi Kata yang Tepat Pada Dokumen Jurnal JTIK Menggunakan Dictionary Lookup dan Damerau-Levenshtein Distance*. 1(6), 498–506.
- Murti, S. (2015). *Eksistensi Penggunaan Bahasa Indonesia di Era Globalisasi*. *Prosiding Seminar Nasional Bulan Bahasa*, 177–184.
- Mustofa. (2019). *Peran Hashtag (#) Dalam Media Sosial A . Pendahuluan Perkembangan teknologi informasi , memungkinkan setiap individu untuk mengubah prioritas kebutuhan hidupnya . Hal ini secara tidak langsung*

berpengaruh terhadap perilaku hidup individu tersebut . Dalam. 7(1), 19–38.

- Madu Muni Adriyani, I Wayan Santiyasa, & Muliantara, A. (2012). *Implementasi Algoritma Levenshtein Distance Dan Metode Empiris Untuk Menampilkan Saran Perbaikan Kesalahan Pengetikan Dokumen Berbahasa Indonesia*. 66, 35–39.
- Pasmita, A. W. (2016). *Analisis Penggunaan Media Sosial Twitter Oleh Pejabat Publik dalam Penerapan Good Governance*. 2–77.
- Putra, M. D. E., Sujaini, H., & Safriadi, N. (2018). *Pengaruh Metode Dictionary Lookup pada Cleaning Korpus Terhadap Akurasi Mesin Penerjemah Statistik Indonesia – Melayu Pontianak*. 4(1), 49–53.
- Robinson, L. (2014). *Implementasi Metode Generalized Vector Space Model Pada Aplikasi Information Retrieval untuk Pencarian Informasi Pada Kumpulan Dokumen Teknik Elektro Di UPT BPI LIPI*. *Jurnal Ilmiah Komputer Dan Informatika (KOMPUTA)*.
- Rehmanawati, Y., & Kusumaningrum, R. (2016). *Studi Perbandingan Algoritma Pencarian String dalam Metode Approximate String Matching untuk Identifikasi Kesalahan Pengetikan Teks*. *Jurnal Buana Informatika*, 7(2), 125–134. <https://doi.org/10.24002/jbi.v7i2.491>
- Rohini, A., & Purnamasari, K. K. (2016). *Pada Hasil Konversi File Dalam Ekstraksi Informasi*.
- Soepomo, P. (2014). *Penerapan Text Mining Pada Sistem Klasifikasi*. *Jurnal Sarjana Teknik Informatika*, 2, 73–83.
- Sun, X., Xu, C., Li, B. I. N., Duan, Y., & Lu, X. (2019). *Enabling Feature Location for API Method Recommendation and Usage Location*. *IEEE Access*, 7, 49872–49881. <https://doi.org/10.1109/ACCESS.2019.2910732>
- Susandi, D., & Sholahudin, U. (2016). *Pemanfaatan Vector Space Model pada Penerapan Algoritma Nazief Adriani , KNN dan Fungsi Similarity Cosine untuk Pembobotan IDF dan WIDF pada Prototipe Sistem Klasifikasi Teks Bahasa Indonesia*. 3(1), 22–29.
- Wangsanegara, N. K., & Subaeki, B. (2015). *Implementasi Natural Language Processing Dalam Pengukuran Ketepatan Ejaan Yang Disempurnakan (Eyd) Pada Abstrak Skripsi Menggunakan Algoritma Fuzzy Logic*. *Jurnal Teknik Informatika*, 8(2) <https://doi.org/10.15408/jti.v8i2.3185>



LAMPIRAN A

DATA HASHTAG DAN HASIL IDENTIFIKASI WORD-BOUNDARY

No	Hashtag Awal	Hasil Identifikasi Word-Boundary	Cek Hasil Kata Pada Kamus	Jumlah Kata Benar	Hasil Kata Validator	Penilaian Validator
1	#hukum_mati_koruptor	hukum mati koruptor	hukum ada, mati ada, koruptor ada,	3 dari 3	hukum mati koruptor	Benar
2	#IndonesiaDiAmbangKebangkrutan	Indonesia Di Ambang Kebangkrutan	indonesia ada, di ada, ambang ada, kebangkrutan ada,	4 dari 4	Indonesia Di Ambang Kebangkrutan	Benar
3	#RakyatBukanSapiPerah	Rakyat Bukan Sapi Perah	rakyat ada, bukan ada, sapi ada, perah ada,	4 dari 4	Rakyat Bukan Sapi Perah	Benar
4	#ILCSalahkanAnies	ILC Salahkan Anies	ilc ada, salahkan ada, anies ada,	3 dari 3	ILC Salahkan Anies	Benar
5	#PapuaMengecamOPM	Papua Mengecam OPM	papua ada, mengecam ada, opm ada,	3 dari 3	Papua Mengecam OPM	Benar
6	#BanjirJakarta2020	Banjir Jakarta 2020	banjir ada, jakarta ada, 2020 ada,	3 dari 3	Banjir Jakarta 2020	Benar
7	#Salam30Hari	Salam 30 Hari	salam ada, 30 ada, hari ada,	3 dari 3	Salam 30 Hari	Benar
8	#4niesDatangIndonesiaSenang	4nies Datang Indonesia Senang	4nies tidak ada, datang ada, indonesia ada, senang ada,	3 dari 4	4nies Datang Indonesia Senang	Benar
9	#4niesRajaKambingHitam	4nies Raja Kambing Hitam	4nies tidak ada, raja ada, kambing ada, hitam ada,	3 dari 4	4nies Raja Kambing Hitam	Benar
10	#VIRUS_CORONA	VIRUS CORONA	virus ada, corona ada,	2 dari 2	VIRUS CORONA	Benar
11	#hari_pahlawan	hari pahlawan	hari ada, pahlawan ada,	2 dari 2	hari pahlawan	Benar
12	#hari_sumpah_pemuda	hari sumpah pemuda	hari ada, sumpah ada,	3 dari 3	hari sumpah pemuda	Benar

			pemuda ada,			
13	#libur14hari	libur 14 hari	libur ada, 14 ada, hari ada,	3 dari 3	libur 14 hari	Benar
14	#terimakasih2019	terimakasih 2019	terimakasih ada, 2019 ada,	2 dari 2	terimakasih 2019	Benar
15	#jokowiMUNDUR	jokowi MUNDUR	jokowi ada, mundur ada,	2 dari 2	jokowi MUNDUR	Benar
16	#SelasaRebahan	Selasa Rebahan	selasa ada, rebahan ada,	2 dari 2	Selasa Rebahan	Benar
17	#KemajuanNyataDiPapua	Kemajuan Nyata Di Papua	kemajuan ada, nyata ada, di ada, papua ada,	4 dari 4	Kemajuan Nyata Di Papua	Benar
18	#IisDahliaNorak	Iis Dahlia Norak	iis ada, dahlia ada, norak ada,	3 dari 3	Iis Dahlia Norak	Benar
19	#TanteMirna	Tante Mirna	tante ada, mirna ada,	2 dari 2	Tante Mirna	Benar
20	#DemoDibayarTapiAmbyar	Demo Dibayar Tapi Ambyar	demo ada, dibayar ada, tapi ada, ambyar ada,	4 dari 4	Demo Dibayar Tapi Ambyar	Benar
21	#TolakPemimpinKafir	Tolak Pemimpin Kafir	tolak ada, pemimpin ada, kafir ada,	3 dari 3	Tolak Pemimpin Kafir	Benar
22	#TolakPemimpinWanita	Tolak Pemimpin Wanita	tolak ada, pemimpin ada, wanita ada,	3 dari 3	Tolak Pemimpin Wanita	Benar
23	#BeritaTerkini	Berita Terkini	berita ada, terkini ada,	2 dari 2	Berita Terkini	Benar
24	#BeritaViral	berita viral	berita ada, viral ada,	2 dari 2	berita viral	Benar
25	#SenyumAja	Senyum Aja	senyum ada, aja ada,	2 dari 2	Senyum Aja	Benar
26	#UntukPapuaMaju	Untuk Papua Maju	untuk ada, papua ada, maju ada,	3 dari 3	Untuk Papua Maju	Benar
27	#BoikotProdukCinaSekarang	Boikot Produk Cina Sekarang	boikot ada, produk ada, cina ada, sekarang ada,	4 dari 4	Boikot Produk Cina Sekarang	Benar
28	#PapuaPastiMakmur	Papua Pasti Makmur	papua ada, pasti ada, makmur ada,	3 dari 3	Papua Pasti Makmur	Benar
29	#lebahganteng	lebah ganteng	lebah ada, ganteng ada,	2 dari 2	lebah ganteng	Benar

30	#bandunggeger	bandung geger	bandung ada, geger ada,	2 dari 2	bandung geger	Benar
31	#hari patah hatinasional	hari patah hati nasi onal	hari ada, patah ada, hati ada, nasi ada, onal tidak ada,	4 dari 5	hari patah hati nasional	Salah
32	#RabuAmbyar	rabu ambyar	rabu ada, ambyar ada,	2 dari 2	rabu ambyar	Benar
33	#LiburanKeJogjaIstimewa	Liburan Ke Jogja Itu Istimewa	liburan ada, ke ada, jogja ada, itu ada, istimewa ada,	5 dari 5	Liburan Ke Jogja Itu Istimewa	Benar
34	#SayaBukanArtis	Saya Bukan Artis	saya ada, bukan ada, artis ada,	3 dari 3	Saya Bukan Artis	Benar
35	#JiwasrayaDiselamatkan	Jiwasraya Diselamatkan	jiwasraya ada, diselamatkan ada,	2 dari 2	Jiwasraya Diselamatka n	Benar
36	#RealisasiResolusi	Realisasi Resolusi	realisasi ada, resolusi ada,	2 dari 2	Realisasi Resolusi	Benar
37	#SalamDamaiIndonesia	salam damai indonesia ku	salam ada, damai ada, indonesia ada, ku ada,	4 dari 4	salam damai indonesia ku	Benar
38	#FilmBucin	Film Bucin	film ada, bucin ada,	2 dari 2	Film Bucin	Benar
39	#ArielTatum	Ariel Tatum	ariel ada, tatum ada,	2 dari 2	Ariel Tatum	Benar
40	#KamisAmbyar	Kamis Ambyar	kamis ada, ambyar ada,	2 dari 2	Kamis Ambyar	Benar
41	#PesanUntukMantan	Pesan Untuk Mantan	pesan ada, untuk ada, mantan ada,	3 dari 3	Pesan Untuk Mantan	Benar
42	#NgajiBawaTumbler	Ngaji Bawa Tumbler	ngaji ada, bawa ada, tumbler ada,	3 dari 3	Ngaji Bawa Tumbler	Benar
43	#MalamJumat	malam jumat	malam ada, jumat ada,	2 dari 2	malam jumat	Benar
44	#JumatBerkah	Jumat Berkah	jumat ada, berkah ada,	2 dari 2	Jumat Berkah	Benar
45	#JumatAmbyar	jumat ambyar	jumat ada, ambyar ada,	2 dari 2	jumat ambyar	Benar
46	#JanganRebahanMulu	jangan rebahan mulu	jangan ada, rebahan ada, mulu ada,	3 dari 3	jangan rebahan mulu	Benar

47	#DukungPansusBanjirDKI	Dukung Pansus Banjir DKI	dukung ada, pansus ada, banjir ada, dki ada,	4 dari 4	Dukung Pansus Banjir DKI	Benar
48	#PapuaKitongMaju	Papua Kitong Maju	papua ada, kitong ada, maju ada,	3 dari 3	Papua Kitong Maju	Benar
49	#DamaiPapua	Damai Papua	damai ada, papua ada,	2 dari 2	Damai Papua	Benar
50	#BersihRumahBanjir	bersih rumah banjir	bersih ada, rumah ada, banjir ada,	3 dari 3	bersih rumah banjir	Benar
51	#MenangkanPancasila	Menangkan Pancasila	menangkan ada, pancasila ada,	2 dari 2	Menangkan Pancasila	Benar
52	#BanjirJakarta	Banjir Jakarta	banjir ada, jakarta ada,	2 dari 2	Banjir Jakarta	Benar
53	#HariJadiTwitterSaya	Hari Jadi Twitter Saya	hari ada, jadi ada, twitter ada, saya ada,	4 dari 4	Hari Jadi Twitter Saya	Benar
54	#HariSenyumNasional	Hari Senyum Nasional	hari ada, senyum ada, nasional ada,	3 dari 3	Hari Senyum Nasional	Benar
55	#TemanTapiMenikah2	Teman Tapi Menikah 2	teman ada, tapi ada, menikah ada, 2 ada,	4 dari 4	Teman Tapi Menikah 2	Benar
56	#TerimakasihMelati	Terimakasih Melati	terimakasih ada, melati ada,	2 dari 2	Terimakasih Melati	Benar
57	#ManaToleransimu	Mana Toleransimu	mana ada, toleransimu ada,	2 dari 2	Mana Toleransimu	Benar
58	#MahasiswaBergerakMelawan	Mahasiswa Bergerak Melawan	mahasiswa ada, bergerak ada, melawan ada,	3 dari 3	Mahasiswa Bergerak Melawan	Benar
59	#VirusCorona	Virus Corona	virus ada, corona ada,	2 dari 2	Virus Corona	Benar
60	#AkhirZaman	Akhir Zaman	akhir ada, zaman ada,	2 dari 2	Akhir Zaman	Benar
61	#AkuCintaKalian	Aku Cinta Kalian	aku ada, cinta ada, kalian ada,	3 dari 3	Aku Cinta Kalian	Benar
62	#MusimHujanBanjirKenangan	Musim Hujan Banjir Kenangan	musim ada, hujan ada, banjir ada, kenangan ada,	4 dari 4	Musim Hujan Banjir Kenangan	Benar

63	#salamdamai	salam damai	salam ada, damai ada,	2 dari 2	salam damai	Benar
64	#AkuBisa	Aku Bisa	aku ada, bisa ada,	2 dari 2	Aku Bisa	Benar
65	#SusuGratisDariNanaz	Susu Gratis Dari Nanaz	susu ada, gratis ada, dari ada, nanaz ada,	4 dari 4	Susu Gratis Dari Nanaz	Benar
66	#Malamminggu	malam minggu	malam ada, minggu ada,	2 dari 2	malam minggu	Benar
67	#BiroJomblo	Biro Jomblo	biro ada, jomblo ada,	2 dari 2	Biro Jomblo	Benar
68	#SabtuRebahan	Sabtu Rebahan	sabtu ada, rebahan ada,	2 dari 2	Sabtu Rebahan	Benar
69	#SabtuCeria	Sabtu Ceria	sabtu ada, ceria ada,	2 dari 2	Sabtu Ceria	Benar
70	#SayangKucingOren	Sayang Kucing Oren	sayang ada, kucing ada, oren ada,	3 dari 3	Sayang Kucing Oren	Benar
71	#SayangLagiApa	Sayang Lagi Apa	sayang ada, lagi ada, apa ada,	3 dari 3	Sayang Lagi Apa	Benar
72	#Selamatpagijakarta	selamat pagi jakarta	selamat ada, pagi ada, jakarta ada,	3 dari 3	selamat pagi jakarta	Benar
73	#Bahagiabersamakamu	bahagia bersama kamu	bahagia ada, bersama ada, kamu ada,	3 dari 3	bahagia bersama kamu	Benar
74	#Akurindu	aku rindu	aku ada, rindu ada,	2 dari 2	aku rindu	Benar
75	#Minggupagi	minggu pagi	minggu ada, pagi ada,	2 dari 2	minggu pagi	Benar
76	#Cuitanrabu	cuitan rabu	cuitan ada, rabu ada,	2 dari 2	cuitan rabu	Benar
77	#Ratukalajengking	ratu kalajengking	ratu ada, kalajengking ada,	2 dari 2	ratu kalajengking	Benar
78	#PuteriKeranda	Puteri Keranda	puteri ada, keranda ada,	2 dari 2	Puteri Keranda	Benar
79	#UntukIndonesia	Untuk Indonesia	untuk ada, indonesia ada,	2 dari 2	Untuk Indonesia	Benar
80	#JanganBersedih	Jangan Bersedih	jangan ada, bersedih ada,	2 dari 2	Jangan Bersedih	Benar
81	#Akumohon	aku mohon	aku ada, mohon ada,	2 dari 2	aku mohon	Benar
82	#Seninsemangat	senin semangat	senin ada, semangat ada,	2 dari 2	senin semangat	Benar

83	#DramaTwitter	Drama Twitter	drama ada, twitter ada,	2 dari 2	Drama Twitter	Benar
84	#bedahrumah	beda hrumah	beda ada, hrumah tidak ada,	1 dari 2	bedah rumah	Salah
85	#dramakorea	drama korea	drama ada, korea ada,	2 dari 2	drama korea	Benar
86	#tidakgalau	tidak galau	tidak ada, galau ada,	2 dari 2	tidak galau	Benar
87	#selasahujan	selasa hujan	selasa ada, hujan ada,	2 dari 2	selasa hujan	Benar
88	#PupukIndonesia	Pupuk Indonesia	pupuk ada, indonesia ada,	2 dari 2	Pupuk Indonesia	Benar
89	#KAMUHARUSTAU	kamu harus tau	kamu ada, harus ada, tau ada,	3 dari 3	kamu harus tau	Benar
90	#calonmenantuidaman	calon menantu idaman	calon ada, menantu ada, idaman ada,	3 dari 3	calon menantu idaman	Benar
91	#JanganLupaMakan	Jangan Lupa Makan	jangan ada, lupa ada, makan ada,	3 dari 3	Jangan Lupa Makan	Benar
92	#HukumTajamKeMuslim	Hukum Tajam Ke Muslim	hukum ada, tajam ada, ke ada, muslim ada,	4 dari 4	Hukum Tajam Ke Muslim	Benar
93	#Hukumberatparakorupt	hukum berat para koruptor	hukum ada, berat ada, para ada, koruptor ada,	4 dari 4	hukum berat para koruptor	Benar
94	#Hukummatikoruptor	hukum mati koruptor	hukum ada, mati ada, koruptor ada,	3 dari 3	hukum mati koruptor	Benar
95	#IslamIdeologiPerjuanganMahasiswa	Islam Ideologi Perjuangan Mahasiswa	islam ada, ideologi ada, perjuangan ada, mahasiswa ada,	4 dari 4	Islam Ideologi Perjuangan Mahasiswa	Benar
96	#BongkarMegaKorupsi	Bongkar Mega Korupsi	bongkar ada, mega ada, korupsi ada,	3 dari 3	Bongkar Mega Korupsi	Benar
97	#RakyatBukanSapiPerah	Rakyat Bukan Sapi Perah	rakyat ada, bukan ada, sapi ada, perah ada,	4 dari 4	Rakyat Bukan Sapi Perah	Benar
98	#RezimSumberGaduh	Rezim Sumber Gaduh	rezim ada, sumber ada, gaduh ada,	3 dari 3	Rezim Sumber Gaduh	Benar
99	#IslamSelamatkanNegeri	Islam Selamatkan	islam ada, selamatkan	3 dari 3	Islam Selamatkan	Benar

		Negeri	ada, negeri ada,		Negeri	
100	#IndonesiaDaruratKorupsi	Indonesia Darurat Korupsi	indonesia ada, darurat ada, korupsi ada,	3 dari 3	Indonesia Darurat Korupsi	Benar
101	#SelamatkanNegeri	Selamatkan Negeri	selamatkan ada, negeri ada,	2 dari 2	Selamatkan Negeri	Benar
102	#RakyatSemakinSengsar	Rakyat Semakin Sengsara	rakyat ada, semakin ada, sengsara ada,	3 dari 3	Rakyat Semakin Sengsara	Benar
103	#MahasiswaTurunKeJalan	Mahasiswa Turun Ke Jalan	mahasiswa ada, turun ada, ke ada, jalan ada,	4 dari 4	Mahasiswa Turun Ke Jalan	Benar
104	#MeroketSkandalKorupsinya	Meroket Skandal Korupsinya	meroket ada, skandal ada, korupsinya ada,	3 dari 3	Meroket Skandal Korupsinya	Benar
105	#JanganBungkamSuaraRakyat	Jangan Bungkam Suara Rakyat	jangan ada, bungkam ada, suara ada, rakyat ada,	4 dari 4	Jangan Bungkam Suara Rakyat	Benar
106	#Janganlupasarapan	jangan lupa sarapan	jangan ada, lupa ada, sarapan ada,	3 dari 3	jangan lupa sarapan	Benar
107	#Wajahbaruindonesia	wajah baru indonesia	wajah ada, baru ada, indonesia ada,	3 dari 3	wajah baru indonesia	Benar
108	#JagaPapuaBersama	Jaga Papua Bersama	jaga ada, papua ada, bersama ada,	3 dari 3	Jaga Papua Bersama	Benar
109	#KentalManisBukanSusBalita	Kental Manis Bukan Susu Balita	kental ada, manis ada, bukan ada, susu ada, balita ada,	5 dari 5	Kental Manis Bukan Susu Balita	Benar
110	#RabuMerindu	Rabu Merindu	rabu ada, merindu ada,	2 dari 2	Rabu Merindu	Benar
111	#MataNajwa	Mata Najwa	mata ada, najwa ada,	2 dari 2	Mata Najwa	Benar
112	#KeahlianTependam	Keahlian Tependam	keahlian ada, tependam ada,	2 dari 2	Keahlian Tependam	Benar
113	#AirBersih	Air Bersih	air ada, bersih ada,	2 dari 2	Air Bersih	Benar
114	#CewekButuhKepastian	Cewek Butuh Kepastian	cewek ada, butuh ada, kepastian ada,	3 dari 3	Cewek Butuh Kepastian	Benar

115	#TokoBarangMantan	Toko Barang Mantan	toko ada, barang ada, mantan ada,	3 dari 3	Toko Barang Mantan	Benar
116	#UdahPutusinAja	Udah Putusin Aja	udah ada, putusin ada, aja ada,	3 dari 3	Udah Putusin Aja	Benar
117	#BerikanArtiLebih	Berikan Arti Lebih	berikan ada, arti ada, lebih ada,	3 dari 3	Berikan Arti Lebih	Benar
118	#BahasJODOH	bahas jodoh	bahas ada, jodoh ada,	2 dari 2	bahas jodoh	Benar
119	#LucintaLuna	Lucinta Luna	lucinta ada, luna ada,	2 dari 2	Lucinta Luna	Benar
120	#MencariJodohDiTwitter	Mencari Jodoh Di Twitter	mencari ada, jodoh ada, di ada, twitter ada,	4 dari 4	Mencari Jodoh Di Twitter	Benar
121	#Jomblobahagia	jomblo bahagia	jomblo ada, bahagia ada,	2 dari 2	jomblo bahagia	Benar
122	#Sinemakeluarga	sinema ke luar ga	sinema ada, ke ada, luar ada, ga tidak ada,	salah	sinema keluarga	Benar
123	#4TahunSekaliUltah	4 Tahun Sekali Ultah	4 ada, tahun ada, sekali ada, ultah ada,	4 dari 4	4 Tahun Sekali Ultah	Benar
124	#Berbagikisah	berbagi kisah	berbagi ada, kisah ada,	2 dari 2	berbagi kisah	Benar
125	#RadionyaPersib	Radionya Persib	radionya ada, persib ada,	2 dari 2	Radionya Persib	Benar
126	#RecehkanTwitter	Recehkan Twitter	recehkan ada, twitter ada,	2 dari 2	Recehkan Twitter	Benar
127	#BesokSenin	Besok Senin	besok ada, senin ada,	2 dari 2	Besok Senin	Benar
128	#CoronaVirusIndonesia	Corona Virus Indonesia	corona ada, virus ada, indonesia ada,	3 dari 3	Corona Virus Indonesia	Benar
129	#CeritaCintaCita	Cerita Cinta Cita	cerita ada, cinta ada, cita ada,	3 dari 3	Cerita Cinta Cita	Benar
130	#Mingguambyar	minggu ambyar	minggu ada, ambyar ada,	2 dari 2	minggu ambyar	Benar
131	#DirumahAjaDulu	Dirumah Aja Dulu	dirumah ada, aja ada, dulu ada,	3 dari 3	Dirumah Aja Dulu	Benar
132	#TolakFCTC	Tolak FCTC	tolak ada, fctc ada,	2 dari 2	Tolak FCTC	Benar

133	#Kamar66Ribu	Kamar 66 Ribu	kamar ada, 66 ada, ribu ada,	3 dari 3	Kamar 66 Ribu	Benar
134	#11Maret	11 Maret	11 ada, maret ada,	2 dari 2	11 Maret	Benar
135	#KamisHalu	Kamis Halu	kamis ada, halu ada,	2 dari 2	Kamis Halu	Benar
136	#1001AlasanGabener	1001 Alasan Gabener	1001 ada, alasan ada, gabener ada,	3 dari 3	1001 Alasan Gabener	Benar
137	#BoikotAice	Boikot Aice	boikot ada, aice ada,	2 dari 2	Boikot Aice	Benar
138	#RumahUya	Rumah Uya	rumah ada, uya ada,	2 dari 2	Rumah Uya	Benar
139	#BerbagiMiniso	Berbagi Miniso	berbagi ada, miniso ada,	2 dari 2	Berbagi Miniso	Benar
140	#TolakKerasOmnibusLaw	Tolak Keras Omnibus Law	tolak ada, keras ada, omnibus ada, law ada,	4 dari 4	Tolak Keras Omnibus Law	Benar
141	#MogokSampaiMenang	Mogok Sampai Menang	mogok ada, sampai ada, menang ada,	3 dari 3	Mogok Sampai Menang	Benar
142	#MarufAminNgapain	Maruf Amin Ngapain	maruf ada, amin ada, ngapain ada,	3 dari 3	Maruf Amin Ngapain	Benar
143	#DukungOmnibusLaw	Dukung Omnibus Law	dukung ada, omnibus ada, law ada,	3 dari 3	Dukung Omnibus Law	Benar
144	#WaspadaVirusCorona	Waspada Virus Corona	waspada ada, virus ada, corona ada,	3 dari 3	Waspada Virus Corona	Benar
145	#MamaBisa	Mama Bisa	mama ada, bisa ada,	2 dari 2	Mama Bisa	Benar
146	#BanserBerduka	Banser Berduka	banser ada, berduka ada,	2 dari 2	Banser Berduka	Benar
147	#TakAdaUangLebaranMelayang	Tak Ada Uang Lebaran Melayang	tak ada, ada ada, uang ada, lebaran ada, melayang ada,	5 dari 5	Tak Ada Uang Lebaran Melayang	Benar
148	#BongkarSelingkuhanSuamiMawarAFI	Bongkar Selingkuhan Suami Mawar AFI	bongkar ada, selingkuhan ada, suami ada, mawar ada, afi ada,	5 dari 5	Bongkar Selingkuhan Suami Mawar AFI	Benar
149	#KKNDiDesaPenari	KKN Di Desa Penari	kkn ada, di ada, desa ada, penari ada,	4 dari 4	KKN Di Desa Penari	Benar
150	#papuamajuberprestasi	papua maju berprestasi	papua ada, maju ada, berprestasi	3 dari 3	papua maju berprestasi	Benar

			ada,			
151	#UsirDubesIndia	Usir Dubes India	usir ada, dubes ada, india ada,	3 dari 3	Usir Dubes India	Benar
152	#AndikaPerkasa	Andika Perkasa	andika ada, perkasa ada,	2 dari 2	Andika Perkasa	Benar
153	#PertaminaBersinergi	pertamina bersinergi	pertamina ada, bersinergi ada,	2 dari 2	pertamina bersinergi	Benar
154	#BangsatBangsa	Bangsat Bangsa	bangsat ada, bangsa ada,	2 dari 2	Bangsat Bangsa	Benar
155	#IndonesiaMajuCumaHal	Indonesia Maju Cuma Halu	indonesia ada, maju ada, cuma ada, halu ada,	4 dari 4	Indonesia Maju Cuma Halu	Benar
156	#NegriNgeri	negri ngeri	negri ada, ngeri ada,	2 dari 2	negri ngeri	Benar
157	#SpiritBaruAnsorJatim	Spirit Baru Ansor Jatim	spirit ada, baru ada, ansor ada, jatim ada,	4 dari 4	Spirit Baru Ansor Jatim	Benar
158	#GemaySantuyTVXQ	Gemay Santuy TVXQ	gemay ada, santuy ada, tvxq ada,	3 dari 3	Gemay Santuy TVXQ	Benar
159	#BauAmiesPolitisiCorona	Bau Amies Politisi Corona	bau ada, amies ada, politisi ada, corona ada,	4 dari 4	Bau Amies Politisi Corona	Benar
160	#FPUJagaPerdamaianDunia	FPU Jaga Perdamaian Dunia	fpu ada, jaga ada, perdamaian ada, dunia ada,	4 dari 4	FPU Jaga Perdamaian Dunia	Benar
161	#StopPolitisasiCorona	Stop Politisasi Corona	stop ada, politisasi ada, corona ada,	3 dari 3	Stop Politisasi Corona	Benar
162	#DiRumahAja	di rumah aja	di ada, rumah ada, aja ada,	3 dari 3	di rumah aja	Benar
163	#MuhasabahAntum	muhasabah antum	muhasabah ada, antum ada,	2 dari 2	muhasabah antum	Benar
164	#PesantrenGratisKlaten	pesan tren gratis klaten	pesan ada, tren ada, gratis ada, klaten ada,	4 dari 4	pesantren gratis klaten	Salah
165	#NolakKeras	Nolak Keras	nolak ada, keras ada,	2 dari 2	Nolak Keras	Benar

166	#transjakarta	trans jakarta	trans ada, jakarta ada,	2 dari 2	trans jakarta	Benar
167	#PercayaKumparan	Percaya Kumparan	percaya ada, kumparan ada,	2 dari 2	Percaya Kumparan	Benar
168	#percaya risalah nabi	percaya risalah nabi	percaya ada, risa ada, lah ada, nabi ada,	4 dari 4	percaya risalah nabi	Salah
169	#azab penimbun masker	azab penimbun masker	azab ada, penimbun ada, masker ada,	3 dari 3	azab penimbun masker	Benar
170	#Rakyat Bersatu Lawan Corona	Rakyat Bersatu Lawan Corona	rakyat ada, bersatu ada, lawan ada, corona ada,	4 dari 4	Rakyat Bersatu Lawan Corona	Benar
171	#Kamis Manis	Kamis Manis	kamis ada, manis ada,	2 dari 2	Kamis Manis	Benar
172	#Bersatu Hadapi Corona	Bersatu Hadapi Corona	bersatu ada, hadapi ada, corona ada,	3 dari 3	Bersatu Hadapi Corona	Benar
173	#Polri Bantu Cegah Corona	Polri Bantu Cegah Corona	polri ada, bantu ada, cegah ada, corona ada,	4 dari 4	Polri Bantu Cegah Corona	Benar
174	#bonus deposit	bonus deposit	bonus ada, deposit ada,	2 dari 2	bonus deposit	Benar
175	#agen judi	agen judi	agen ada, judi ada,	2 dari 2	agen judi	Benar
176	#Stop Sementara Ekspor Masker	Stop Sementara Ekspor Masker	stop ada, sementara ada, ekspor ada, masker ada,	4 dari 4	Stop Sementara Ekspor Masker	Benar
177	#masker mahal	masker mahal	masker ada, mahal ada,	2 dari 2	masker mahal	Benar
178	#BPJSEdukasi Corona	BPJS Edukasi Corona	bpjs ada, edukasi ada, corona ada,	3 dari 3	BPJS Edukasi Corona	Benar
179	#Govinda Hal Hebat	Govinda Hal Hebat	govinda ada, hal ada, hebat ada,	3 dari 3	Govinda Hal Hebat	Benar
180	#perusahaan pengiriman barang	perusahaan pengiriman barang	perusahaan ada, pengiriman ada, barang ada,	3 dari 3	perusahaan pengiriman barang	Benar
181	#Arab Saudi	Arab Saudi	arab ada, saudi ada,	2 dari 2	Arab Saudi	Benar

182	#CoronaBukaBorokRezi	Corona Buka Borok Rezim	corona ada, buka ada, borok ada, rezim ada,	4 dari 4	Corona Buka Borok Rezim	Benar
183	#BUMNLawanCorona	BUMN Lawan Corona	bumn ada, lawan ada, corona ada,	3 dari 3	BUMN Lawan Corona	Benar
184	#KamiTidakTakutVirusCorona	Kami Tidak Takut Virus Corona	kami ada, tidak ada, takut ada, virus ada, corona ada,	5 dari 5	Kami Tidak Takut Virus Corona	Benar
185	#StopPanikLawanCorona	Stop Panik Lawan Corona	stop ada, panik ada, lawan ada, corona ada,	4 dari 4	Stop Panik Lawan Corona	Benar
186	#Bersihbersihrumah	bersih bersih rumah	bersih ada, bersih ada, rumah ada,	3 dari 3	bersih bersih rumah	Benar
187	#Jualsouvenirbantalmobilpromosi	jual souvenir bantal mobil promosi	jual ada, souvenir ada, bantal ada, mobil ada, promosi ada,	5 dari 5	jual souvenir bantal mobil promosi	Benar
188	#AzabPenimbunGula	Azab Penimbun Gula	azab ada, penimbun ada, gula ada,	3 dari 3	Azab Penimbun Gula	Benar
189	#AzabGabenerBodong	Azab Gabener Bodong	azab ada, gabener ada, bodong ada,	3 dari 3	Azab Gabener Bodong	Benar
190	#NegeriGaduhUlahPenguasa	Negeri Gaduh Ulah Penguasa	negeri ada, gaduh ada, ulah ada, penguasa ada,	4 dari 4	Negeri Gaduh Ulah Penguasa	Benar
191	#PecatSittiHikmawatty	Pecat Sitti Hikmawatty	pecat ada, sitti ada, hikmawatty ada,	3 dari 3	Pecat Sitti Hikmawatty	Benar
192	#WajahBanjirJakarta	Wajah Banjir Jakarta	wajah ada, banjir ada, jakarta ada,	3 dari 3	Wajah Banjir Jakarta	Benar
193	#IndonesiaTanpaPacaran	Indonesia Tanpa Pacaran	indonesia ada, tanpa ada, pacaran ada,	3 dari 3	Indonesia Tanpa Pacaran	Benar
194	#Coronadisekitarkita	corona di sekitar kita	corona ada, di ada, sekitar ada, kita ada,	4 dari 4	corona di sekitar kita	Benar
195	#DisiniBanjirKamuDimana	Disini Banjir Kamu Dimana	disini ada, banjir ada, kamu ada, dimana ada,	4 dari 4	Disini Banjir Kamu Dimana	Benar
196	#hari_ibu	hari ibu	hari ada, ibu	2 dari 2	hari ibu	Benar

			ada,			
197	#hari_santri_nasional	hari santri nasional	hari ada, santri ada, nasional ada,	3 dari 3	hari santri nasional	Benar
198	#hari_guru	hari guru	hari ada, guru ada,	2 dari 2	hari guru	Benar
199	#hari_kartini	hari kartini	hari ada, kartini ada,	2 dari 2	hari kartini	Benar
200	#hari_raya	hari raya	hari ada, raya ada,	2 dari 2	hari raya	Benar
201	#hari_batik_nasional	hari batik nasional	hari ada, batik ada, nasional ada,	3 dari 3	hari batik nasional	Benar
202	#hari_ini	hari ini	hari ada, ini ada,	2 dari 2	hari ini	Benar
203	#hari_pertama	hari pertama	hari ada, pertama ada,	2 dari 2	hari pertama	Benar
204	#hari_lahir_pancasila	hari lahir pancasila	hari ada, lahir ada, pancasila ada,	3 dari 3	hari lahir pancasila	Benar
205	#hari_pendidikan_nasio	hari pendidikan nasional	hari ada, pendidikan ada, nasional ada,	3 dari 3	hari pendidikan nasional	Benar
206	#hari_jadi	hari jadi	hari ada, jadi ada,	2 dari 2	hari jadi	Benar
207	#hari_anak_nasional	hari anak nasional	hari ada, anak ada, nasional ada,	3 dari 3	hari anak nasional	Benar
208	#hari_bahagia	hari bahagia	hari ada, bahagia ada,	2 dari 2	hari bahagia	Benar
209	#hari_kemenangan	hari kemenangan	hari ada, kemenangan ada,	2 dari 2	hari kemenangan	Benar
210	#hari_kemerdekaan	hari kemerdekaan	hari ada, kemerdekaan ada,	2 dari 2	hari kemerdekaan	Benar
211	#hari_terakhir	hari terakhir	hari ada, terakhir ada,	2 dari 2	hari terakhir	Benar
212	#hari_kesaktian_pancasila	hari kesaktian pancasila	hari ada, kesaktian ada, pancasila ada,	3 dari 3	hari kesaktian pancasila	Benar
213	#hari_ayah	hari ayah	hari ada, ayah ada,	2 dari 2	hari ayah	Benar
214	#hari_pancasila	hari pancasila	hari ada, pancasila ada,	2 dari 2	hari pancasila	Benar
215	#azabkubur	azab ku bur	azab ada, ku ada, bur ada,	3 dari 3	azab kubur	Salah

216	#azabneraka	azab neraka	azab ada, neraka ada,	2 dari 2	azab neraka	Benar
217	#sinetronazab	sinetron azab	sinetron ada, azab ada,	2 dari 2	sinetron azab	Benar
218	#kisahazab	kisah azab	kisah ada, azab ada,	2 dari 2	kisah azab	Benar
219	#janganpernahberubah	jangan pernah berubah	jangan ada, pernah ada, berubah ada,	3 dari 3	jangan pernah berubah	Benar
220	#janganmenyerah	jangan menyerah	jangan ada, menyerah ada,	2 dari 2	jangan menyerah	Benar
221	#janganlupabahagia	jangan lupa bahagia	jangan ada, lupa ada, bahagia ada,	3 dari 3	jangan lupa bahagia	Benar
222	#janganmarah	jangan marah	jangan ada, marah ada,	2 dari 2	jangan marah	Benar
223	#jangankecamsaya	jangan ke cam saya	jangan ada, ke ada, cam ada, saya ada,	4 dari 4	jangan kecam saya	Salah
224	#janganlupabersyukur	jangan lupa bersyukur	jangan ada, lupa ada, bersyukur ada,	3 dari 3	jangan lupa bersyukur	Benar
225	#jangandengerin	jangan dengerin	jangan ada, dengerin ada,	2 dari 2	jangan dengerin	Benar
226	#SelamatDatangDiTwitter	Selamat Datang Di Twitter	selamat ada, datang ada, di ada, twitter ada,	4 dari 4	Selamat Datang Di Twitter	Benar
227	#IbuSudahBisa	Ibu Sudah Bisa	ibu ada, sudah ada, bisa ada,	3 dari 3	Ibu Sudah Bisa	Benar
228	#IbuSampeKaget	Ibu Sampe Kaget	ibu ada, sampe ada, kaget ada,	3 dari 3	Ibu Sampe Kaget	Benar
229	#RecehkanAgustus	Recehkan Agustus	recehkan ada, agustus ada,	2 dari 2	Recehkan Agustus	Benar
230	#AjaibnyaOrangIndonesia	Ajaibnya Orang Indonesia	ajaibnya ada, orang ada, indonesia ada,	3 dari 3	Ajaibnya Orang Indonesia	Benar
231	#MulutNetizenIndonesia	Mulut Netizen Indonesia	mulut ada, netizen ada, indonesia ada,	3 dari 3	Mulut Netizen Indonesia	Benar
232	#gubernurterbodoh	gubernur terbodoh	gubernur ada, terbodoh ada,	2 dari 2	gubernur terbodoh	Benar
233	#KamiWaspadaCorona	Kami Waspada Corona	kami ada, waspada ada, corona ada,	3 dari 3	Kami Waspada Corona	Benar

234	#papamudabisa	papa muda bisa	papa ada, muda ada, bisa ada,	3 dari 3	papa muda bisa	Benar
235	#kasuspanik	kasus panik	kasus ada, panik ada,	2 dari 2	kasus panik	Benar
236	#RakyatInginPerubahan	Rakyat Ingin Perubahan	rakyat ada, ingin ada, perubahan ada,	3 dari 3	Rakyat Ingin Perubahan	Benar
237	#GueNgantor	Gue Ngantor	gue ada, ngantor ada,	2 dari 2	Gue Ngantor	Benar
238	#SayaPercayaJokowi	Saya Percaya Jokowi	saya ada, percaya ada, jokowi ada,	3 dari 3	Saya Percaya Jokowi	Benar
239	#IndonesiaBisa	Indonesia Bisa	indonesia ada, bisa ada,	2 dari 2	Indonesia Bisa	Benar
240	#ruangguru	ruangguru	ruangguru ada,	1 dari 1	ruangguru	Benar
241	#telkomselruangguru	telkomsel ruangguru	telkomsel ada, ruangguru ada,	2 dari 2	telkomsel ruangguru	Benar
242	#jadijuara	jadi juara	jadi ada, juara ada,	2 dari 2	jadi juara	Benar
243	#libur14hari	libur 14 hari	libur ada, 14 ada, hari ada,	3 dari 3	libur 14 hari	Benar
244	#libur2minggu	libur 2 minggu	libur ada, 2 ada, minggu ada,	3 dari 3	libur 2 minggu	Benar
245	#TempatHoror	Tempat Horor	tempat ada, horor ada,	2 dari 2	Tempat Horor	Benar
246	#planetgeografi	planet geografi	planet ada, geografi ada,	2 dari 2	planet geografi	Benar
247	#zebraindonesia	zebra indonesia	zebra ada, indonesia ada,	2 dari 2	zebra indonesia	Benar
248	#besokselasa	besok selasa	besok ada, selasa ada,	2 dari 2	besok selasa	Benar
249	#AsalTauRahasiannya	Asal Tau Rahasiannya	asal ada, tau ada, rahasianya ada,	3 dari 3	Asal Tau Rahasiannya	Benar
250	#gurubesar	guru besar	guru ada, besar ada,	2 dari 2	guru besar	Benar
251	#gurumuda	guru muda	guru ada, muda ada,	2 dari 2	guru muda	Benar
252	#guruhonorer	guru honorer	guru ada, honorer ada,	2 dari 2	guru honorer	Benar
253	#matematikamenyenangkan	matematika menyenangkan	matematika ada, menyenangkan	2 dari 2	matematika menyenangkan	Benar

			ada,			
254	#matematikasusah	matematika susah	matematika ada, susah ada,	2 dari 2	matematika susah	Benar
255	#cara menghitung	cara menghitung	cara ada, menghitung ada,	2 dari 2	cara menghitung	Benar
256	#suka berhitung	suka berhitung	suka ada, berhitung ada,	2 dari 2	suka berhitung	Benar
257	#sekolahpasar modal	sekolah pasar modal	sekolah ada, pasar ada, modal ada,	3 dari 3	sekolah pasar modal	Benar
258	#PartaiSodaritasIndonesia	Partai Sodaritas Indonesia	partai ada, sodaritas ada, indonesia ada,	3 dari 3	Partai Sodaritas Indonesia	Benar
259	#StafKhususPresiden	Staf Khusus Presiden	staf ada, khusus ada, presiden ada,	3 dari 3	Staf Khusus Presiden	Benar
260	#RajaKreatif	Raja Kreatif	raja ada, kreatif ada,	2 dari 2	Raja Kreatif	Benar
261	#HariGuruNasional	Hari Guru Nasional	hari ada, guru ada, nasional ada,	3 dari 3	Hari Guru Nasional	Benar
262	#HariGuru2019	Hari Guru 2019	hari ada, guru ada, 2019 ada,	3 dari 3	Hari Guru 2019	Benar
263	#KodeDiskon	kode di skon	kode ada, di ada, skon tidak ada,	2 dari 3	kode diskon	Salah
264	#Belajarsangatmenyenangkan	belajar sangat menyenangkan	belajar ada, sangat ada, menyenangkan ada,	3 dari 3	belajar sangat menyenangkan	Benar
265	#KebebasanSiswa	Kebebasan Siswa	kebebasan ada, siswa ada,	2 dari 2	Kebebasan Siswa	Benar
266	#Tempatmenulisilmiah	tempat menulis ilmiah	tempat ada, menulis ada, ilmiah ada,	3 dari 3	tempat menulis ilmiah	Benar
267	#Pelatihanmenulis	pelatihan menulis	pelatihan ada, menulis ada,	2 dari 2	pelatihan menulis	Benar
268	#Mejakerja	meja ke rja	meja ada, ke ada, rja tidak ada,	2 dari 3	meja kerja	Salah
269	#Ruangguru5tahun	ruangguru 5 tahun	ruangguru ada, 5 ada, tahun ada,	3 dari 3	ruangguru 5 tahun	Benar
270	#Pastibisa	pasti bisa	pasti ada, bisa ada,	2 dari 2	pasti bisa	Benar

271	#MadrasahHebat	Madrasah Hebat	madrasah ada, hebat ada,	2 dari 2	Madrasah Hebat	Benar
272	#GuruMadrasah	Guru Madrasah	guru ada, madrasah ada,	2 dari 2	Guru Madrasah	Benar
273	#GerakSerentak	Gerak Serentak	gerak ada, serentak ada,	2 dari 2	Gerak Serentak	Benar
274	#SehatiMelayani	Sehati Melayani	sehati ada, melayani ada,	2 dari 2	Sehati Melayani	Benar
275	#KisahInspiratif	Kisah Inspiratif	kisah ada, inspiratif ada,	2 dari 2	Kisah Inspiratif	Benar
276	#ToleransiIndonesia	Toleransi Indonesia	toleransi ada, indonesia ada,	2 dari 2	Toleransi Indonesia	Benar
277	#BedaItuKeren	Beda Itu Keren	beda ada, itu ada, keren ada,	3 dari 3	Beda Itu Keren	Benar
278	#MeyakiniMenghargai	Meyakini Menghargai	meyakini ada, menghargai ada,	2 dari 2	Meyakini Menghargai	Benar
279	#WaspadaPencurian	Waspada Pencurian	waspada ada, pencurian ada,	2 dari 2	Waspada Pencurian	Benar
280	#BelajarJadiLuarBiasa	Belajar Jadi Luar Biasa	belajar ada, jadi ada, luar ada, biasa ada,	4 dari 4	Belajar Jadi Luar Biasa	Benar
281	#BanggaJadiGuru	Bangga Jadi Guru	bangga ada, jadi ada, guru ada,	3 dari 3	Bangga Jadi Guru	Benar
282	#TempatLiburan	Tempat Liburan	tempat ada, liburan ada,	2 dari 2	Tempat Liburan	Benar
283	#WaspadaCegahCorona	Waspada Cegah Corona	waspada ada, cegah ada, corona ada,	3 dari 3	Waspada Cegah Corona	Benar
284	#TolongMenolong	Tolong Menolong	tolong ada, menolong ada,	2 dari 2	Tolong Menolong	Benar
285	#Liburcorona	libur corona	libur ada, corona ada,	2 dari 2	libur corona	Benar
286	#SeninRebahan	Senin Rebahan	senin ada, rebahan ada,	2 dari 2	Senin Rebahan	Benar
287	#KapanNikah	Kapan Nikah	kapan ada, nikah ada,	2 dari 2	Kapan Nikah	Benar
288	#Dirumahsaja	di rumah saja	di ada, rumah ada, saja ada,	3 dari 3	di rumah saja	Benar
289	#Belajardirumah	belajar di rumah	belajar ada, di ada, rumah ada,	3 dari 3	belajar di rumah	Benar

290	#liburterusbosan	libur terus bosan	libur ada, terus ada, bosan ada,	3 dari 3	libur terus bosan	Benar
291	#liburmantapjiwa	libur mantap jiwa	libur ada, mantap ada, jiwa ada,	3 dari 3	libur mantap jiwa	Benar
292	#malesbelajar	males belajar	males ada, belajar ada,	2 dari 2	males belajar	Benar
293	#sinetronindonesia	sinetron indonesia	sinetron ada, indonesia ada,	2 dari 2	sinetron indonesia	Benar
294	#terimakasihcorona	terima kasih corona	terima ada, kasih ada, corona ada,	3 dari 3	terima kasih corona	Benar
295	#LawanCorona	Lawan Corona	lawan ada, corona ada,	2 dari 2	Lawan Corona	Benar
296	#JatengLawanCorona	Jateng Lawan Corona	jateng ada, lawan ada, corona ada,	3 dari 3	Jateng Lawan Corona	Benar
297	#pikirkaryawan	pikiran karyawan	pikiran ada, karyawan ada,	2 dari 2	pikiran karyawan	Benar
298	#Irihati	iri hati	iri ada, hati ada,	2 dari 2	iri hati	Benar
299	#Karyawanirihati	karyawan iri hati	karyawan ada, iri ada, hati ada,	3 dari 3	karyawan iri hati	Benar
300	#MenujuMasjid	Menuju Masjid	menuju ada, masjid ada,	2 dari 2	Menuju Masjid	Benar
301	#MembudayakanSujud	Membudayakan Sujud	membudayakan ada, sujud ada,	2 dari 2	Membudayakan Sujud	Benar
302	#50TahunAnakRevolusi	50 Tahun Anak Revolusi	50 ada, tahun ada, anak ada, revolusi ada,	4 dari 4	50 Tahun Anak Revolusi	Benar
303	#SetorFoto	Setor Foto	setor ada, foto ada,	2 dari 2	Setor Foto	Benar
304	#DukungJokowiBasmiCorona	Dukung Jokowi Basmi Corona	dukung ada, jokowi ada, basmi ada, corona ada,	4 dari 4	Dukung Jokowi Basmi Corona	Benar
305	#JawaBarat	Jawa Barat	jawa ada, barat ada,	2 dari 2	Jawa Barat	Benar
306	#DiamDiRumahCuk	Diam Di Rumah Cuk	diam ada, di ada, rumah ada, cuk ada,	4 dari 4	Diam Di Rumah Cuk	Benar
307	#TergagapCorona	Tergagap Corona	tergagap ada, corona ada,	2 dari 2	Tergagap Corona	Benar
308	#sekolahlibur	sekolah libur	sekolah ada, libur ada,	2 dari 2	sekolah libur	Benar

309	#mataminus	mata minus	mata ada, minus ada,	2 dari 2	mata minus	Benar
310	#solusimataminus	solusi mata minus	solusi ada, mata ada, minus ada,	3 dari 3	solusi mata minus	Benar
311	#solusimatasilinder	solusi mata silinder	solusi ada, mata ada, silinder ada,	3 dari 3	solusi mata silinder	Benar
312	#tipsmatasehat	tips mata sehat	tips ada, mata ada, sehat ada,	3 dari 3	tips mata sehat	Benar
313	#senen senang	senen senang	senen ada, senang ada,	2 dari 2	senen senang	Benar
314	#KuliahOnline	Kuliah Online	kuliah ada, online ada,	2 dari 2	Kuliah Online	Benar
315	#SebelumAkuMengenall inter	Sebelum Aku Mengenal Inter	sebelum ada, aku ada, mengenal ada, inter ada,	4 dari 4	Sebelum Aku Mengenal Inter	Benar
316	#bosanbelajaronline	bosan belajar online	bosan ada, belajar ada, online ada,	3 dari 3	bosan belajar online	Benar
317	#JurnalKomunikator	jurnal komunikator	jurnal ada, komunikator ada,	2 dari 2	jurnal komunikator	Benar
318	#nikahonline	nikah online	nikah ada, online ada,	2 dari 2	nikah online	Benar
319	#hilangkanperutbuncit	hilang kan perut buncit	hilang ada, kan ada, perut ada, buncit ada,	4 dari 4	hilang kan perut buncit	Benar
320	#gagagacorona	gara gara corona	gara ada, gara ada, corona ada,	3 dari 3	gara gara corona	Benar
321	#AnakPeternakanKuliah Online	Anak Peternakan Kuliah Online	anak ada, peternakan ada, kuliah ada, online ada,	4 dari 4	Anak Peternakan Kuliah Online	Benar
322	#KenanganPakeBBM	Kenangan Pake BBM	kenangan ada, pake ada, bbm ada,	3 dari 3	Kenangan Pake BBM	Benar
323	#DahMalasLiatnya	Dah Malas Liatnya	dah ada, malas ada, liatnya ada,	3 dari 3	Dah Malas Liatnya	Benar
324	#KarenaTwitter	Karena Twitter	karena ada, twitter ada,	2 dari 2	Karena Twitter	Benar
325	#HariGiniJanjiDoang	Hari Gini Janji Doang	hari ada, gini ada, janji ada, doang ada,	4 dari 4	Hari Gini Janji Doang	Benar

326	#BuktiCapresMana	Bukti Capres Mana	bukti ada, capres ada, mana ada,	3 dari 3	Bukti Capres Mana	Benar
327	#MantanItu	Mantan Itu	mantan ada, itu ada,	2 dari 2	Mantan Itu	Benar
328	#IndonesiaBersyukur	Indonesia Bersyukur	indonesia ada, bersyukur ada,	2 dari 2	Indonesia Bersyukur	Benar
329	#RinduTelukPanji	rindu teluk panji	rindu ada, teluk ada, panji ada,	3 dari 3	rindu teluk panji	Benar
330	#FACup	FA Cup	fa ada, cup ada,	2 dari 2	FA Cup	Benar
331	#JagaNilaiLuhurBangsa	Jaga Nilai Luhur Bangsa	jaga ada, nilai ada, luhur ada, bangsa ada,	4 dari 4	Jaga Nilai Luhur Bangsa	Benar
332	#SehatMulaiDariKita	Sehat Mulai Dari Kita	sehat ada, mulai ada, dari ada, kita ada,	4 dari 4	Sehat Mulai Dari Kita	Benar
333	#AsikinHidup	Asikin Hidup	asikin ada, hidup ada,	2 dari 2	Asikin Hidup	Benar
334	#SaatYangTerindah	saat yang terindah	saat ada, yang ada, terindah ada,	3 dari 3	saat yang terindah	Benar
335	#HaramPilihPemimpinIngkarJanji	Haram Pilih Pemimpin Ingkar Janji	haram ada, pilih ada, pemimpin ada, ingkar ada, janji ada,	5 dari 5	Haram Pilih Pemimpin Ingkar Janji	Benar
336	#UdahKhilafahAja	Udah Khilafah Aja	udah ada, khilafah ada, aja ada,	3 dari 3	Udah Khilafah Aja	Benar
337	#TernyataSolusinyaKhilafah	Ternyata Solusinya Khilafah	ternyata ada, solusinya ada, khilafah ada,	3 dari 3	Ternyata Solusinya Khilafah	Benar
338	#HaramPilihPemimpinDzalim	Haram Pilih Pemimpin Dzalim	haram ada, pilih ada, pemimpin ada, dzalim ada,	4 dari 4	Haram Pilih Pemimpin Dzalim	Benar
339	#DebatPilpres2019	Debat Pilpres 2019	debat ada, pilpres ada, 2019 ada,	3 dari 3	Debat Pilpres 2019	Benar
340	#PrabowoMenangDebat	Prabowo Menang Debat	prabowo ada, menang ada, debat ada,	3 dari 3	Prabowo Menang Debat	Benar
341	#DebatPintarJokowi	Debat Pintar Jokowi	debat ada, pintar ada, jokowi ada,	3 dari 3	Debat Pintar Jokowi	Benar

342	#DebatKeduaPilpres2019	Debat Kedua Pilpres 2019	debat ada, kedua ada, pilpres ada, 2019 ada,	4 dari 4	Debat Kedua Pilpres 2019	Benar
343	#DebatSebel	Debat Sebel	debat ada, sebel ada,	2 dari 2	Debat Sebel	Benar
344	#AaGymPilihPrabowo	Aa Gym Pilih Prabowo	aa ada, gym ada, pilih ada, prabowo ada,	4 dari 4	Aa Gym Pilih Prabowo	Benar
345	#JokowiMenangTotalDebat	Jokowi Menang Total Debat	jokowi ada, menang ada, total ada, debat ada,	4 dari 4	Jokowi Menang Total Debat	Benar
346	#SiapapunPresidennya	Siapapun Presidennya	siapapun ada, presidennya ada,	2 dari 2	Siapapun Presidennya	Benar
347	#BersatuDemiNegeri	Bersatu Demi Negeri	bersatu ada, demi ada, negeri ada,	3 dari 3	Bersatu Demi Negeri	Benar
348	#JokowiDijebak	Jokowi Dijebak	jokowi ada, dijawab ada,	2 dari 2	Jokowi Dijebak	Benar
349	#JokowiMajuIndonesiaMaju	Jokowi Maju Indonesia Maju	jokowi ada, maju ada, indonesia ada, maju ada,	4 dari 4	Jokowi Maju Indonesia Maju	Benar
350	#TanyaGundala	Tanya Gundala	tanya ada, gundala ada,	2 dari 2	Tanya Gundala	Benar
351	#ProfAndalanPilihanku	Prof Andalan Pilihanku	prof ada, andalan ada, pilihanku ada,	3 dari 3	Prof Andalan Pilihanku	Benar
352	#RiauDibakarBukanTerbakar	Riau Dibakar Bukan Terbakar	riau ada, dibakar ada, bukan ada, terbakar ada,	4 dari 4	Riau Dibakar Bukan Terbakar	Benar
353	#KakolaEnaknyaTerlalu	Kakola Enaknya Terlalu	kakola ada, enaknya ada, terlalu ada,	3 dari 3	Kakola Enaknya Terlalu	Benar
354	#IndonesiaMemilih	Indonesia Memilih	indonesia ada, memilih ada,	2 dari 2	Indonesia Memilih	Benar
355	#KawalHasilSuaraTPS	Kawal Hasil Suara TPS	kawal ada, hasil ada, suara ada, tps ada,	4 dari 4	Kawal Hasil Suara TPS	Benar
356	#Pemilu2019	Pemilu 2019	pemilu ada, 2019 ada,	2 dari 2	Pemilu 2019	Benar
357	#milihjokowijuga	milih jokowi juga	milih ada, jokowi ada, juga ada,	3 dari 3	milih jokowi juga	Benar

358	#salam2jari	salam 2 jari	salam ada, 2 ada, jari ada,	3 dari 3	salam 2 jari	Benar
359	#jokowi9juli	jokowi 9 juli	jokowi ada, 9 ada, juli ada,	3 dari 3	jokowi 9 juli	Benar
360	#presidenjokowi	presiden jokowi	presiden ada, jokowi ada,	2 dari 2	presiden jokowi	Benar
361	#TampilanBaruTwitter	Tampilan Baru Twitter	tampilan ada, baru ada, twitter ada,	3 dari 3	Tampilan Baru Twitter	Benar
362	#InfoKaget	Info Kaget	info ada, kaget ada,	2 dari 2	Info Kaget	Benar
363	#RockyGerungMenghinaPresiden	Rocky Gerung Menghina Presiden	rocky ada, gerung ada, menghina ada, presiden ada,	4 dari 4	Rocky Gerung Menghina Presiden	Benar
364	#GerakTakTerbatas	Gerak Tak Terbatas	gerak ada, tak ada, terbatas ada,	3 dari 3	Gerak Tak Terbatas	Benar
365	#PidatoDuluLengserKemudian	Pidato Dulu Lengser Kemudian	pidato ada, dulu ada, lengser ada, kemudian ada,	4 dari 4	Pidato Dulu Lengser Kemudian	Benar
366	#AhokBikinRusuhNegri	Ahok Bikin Rusuh Negri	ahok ada, bikin ada, rusuh ada, negri ada,	4 dari 4	Ahok Bikin Rusuh Negri	Benar
367	#OrangKayaGakTau	Orang Kaya Gak Tau	orang ada, kaya ada, gak ada, tau ada,	4 dari 4	Orang Kaya Gak Tau	Benar
368	#KepungKedubesChina	Kepung Kedubes China	kepung ada, kedubes ada, china ada,	3 dari 3	Kepung Kedubes China	Benar
369	#KamiBersatuBersamaUAS	Kami Bersatu Bersama UAS	kami ada, bersatu ada, bersama ada, uas ada,	4 dari 4	Kami Bersatu Bersama UAS	Benar
370	#BatalkanPindahIbuKota	Batalkan Pindah Ibu Kota	batalkan ada, pindah ada, ibu ada, kota ada,	4 dari 4	Batalkan Pindah Ibu Kota	Benar
371	#YangGajiKamuSiapa	Yang Gaji Kamu Siapa	yang ada, gaji ada, kamu ada, siapa ada,	4 dari 4	Yang Gaji Kamu Siapa	Benar
372	#KamuDigajiSiapa	Kamu Digaji Siapa	kamu ada, digaji ada, siapa ada,	3 dari 3	Kamu Digaji Siapa	Benar
373	#InfrastrukturLangitnyaKyai	Infrastruktur Langitnya Kyai	infrastruktur ada, langitnya ada, kyai ada,	3 dari 3	Infrastruktur Langitnya Kyai	Benar

374	#DoakanPrabowoSandi	Doakan Prabowo Sandi	doakan ada, prabowo ada, sandi ada,	3 dari 3	Doakan Prabowo Sandi	Benar
375	#SaksiTanpaDrama	Saksi Tanpa Drama	saksi ada, tanpa ada, drama ada,	3 dari 3	Saksi Tanpa Drama	Benar
376	#RakyatTolakHasilPilpres	Rakyat Tolak Hasil Pilpres	rakyat ada, tolak ada, hasil ada, pilpres ada,	4 dari 4	Rakyat Tolak Hasil Pilpres	Benar
377	#JokowiAminSudahMenang	Jokowi Amin Sudah Menang	jokowi ada, amin ada, sudah ada, menang ada,	4 dari 4	Jokowi Amin Sudah Menang	Benar
378	#JagaMantanJagaKenang	Jaga Mantan Jaga Kenangan	jaga ada, mantan ada, jaga ada, kenangan ada,	4 dari 4	Jaga Mantan Jaga Kenangan	Benar
379	#GebyarSamosirSigalegal	Gebyar Samosir Sigalegale 2019	gebyar ada, samosir ada, sigalegale ada, 2019 ada,	4 dari 4	Gebyar Samosir Sigalegale 2019	Benar
380	#BanserJagaBumi	Banser Jaga Bumi	banser ada, jaga ada, bumi ada,	3 dari 3	Banser Jaga Bumi	Benar
381	#BundaDigital	Bunda Digital	bunda ada, digital ada,	2 dari 2	Bunda Digital	Benar
382	#KPUJanganSalahInputData	KPU Jangan Salah Input Data	kpu ada, jangan ada, salah ada, input ada, data ada,	5 dari 5	KPU Jangan Salah Input Data	Benar
383	#HariBukuSedunia	Hari Buku Sedunia	hari ada, buku ada, sedunia ada,	3 dari 3	Hari Buku Sedunia	Benar
384	#KemarauProduktif	Kemarau Produktif	kemarau ada, produktif ada,	2 dari 2	Kemarau Produktif	Benar
385	#PerempuanHebat	perempuan hebat	perempuan ada, hebat ada,	2 dari 2	perempuan hebat	Benar
386	#LayangAnputus	layang anputus	layang ada, anputus tidak ada,	1 dari 2	layangan putus	Salah
387	#RiniPerampokNegeri	Rini Perampok Negeri	rini ada, perampok ada, negeri ada,	3 dari 3	Rini Perampok Negeri	Benar
388	#HariKebangkitanNasional	Hari Kebangkitan Nasional	hari ada, kebangkitan ada, nasional ada,	3 dari 3	Hari Kebangkitan Nasional	Benar

389	#JokowiSukaKerja2024	Jokowi Suka Kerja 2024	jokowi ada, suka ada, kerja ada, 2024 ada,	4 dari 4	Jokowi Suka Kerja 2024	Benar
390	#golektakjil	Golek takjil	golek ada, takjil ada,	2 dari 2	golek takjil	Benar
391	#SabtuMingguSiaran	Sabtu Minggu Siaran	sabtu ada, minggu ada, siaran ada,	3 dari 3	Sabtu Minggu Siaran	Benar
392	#WajahmuPlastik	Wajahmu Plastik	wajahmu ada, plastik ada,	2 dari 2	Wajahmu Plastik	Benar
393	#KoalisiPlastik	Koalisi Plastik	koalisi ada, plastik ada,	2 dari 2	Koalisi Plastik	Benar
394	#MisteriHilangnyaSandiagUno	Misteri Hilangnya Sandiaga Uno	misteri ada, hilangnya ada, sandiaga ada, uno ada,	4 dari 4	Misteri Hilangnya Sandiaga Uno	Benar
395	#BolosHariKejepit	Bolos Hari Kejepit	bolos ada, hari ada, kejepit ada,	3 dari 3	Bolos Hari Kejepit	Benar
396	#InginKuBerkataKasar	Ingin Ku Berkata Kasar	ingin ada, ku ada, berkata ada, kasar ada,	4 dari 4	Ingin Ku Berkata Kasar	Benar
397	#PrabowoHalu	Prabowo Halu	prabowo ada, halu ada,	2 dari 2	Prabowo Halu	Benar
398	#HumorRecehKu	Humor Receh Ku	humor ada, receh ada, ku ada,	3 dari 3	Humor Receh Ku	Benar
399	#HariKecepatNasional	Hari Kecepat Nasional	hari ada, kecepatan ada, nasional ada,	3 dari 3	Hari Kecepat Nasional	Benar
400	#UdahDong	Udah Dong	udah ada, dong ada,	2 dari 2	Udah Dong	Benar
401	#Saya01	Saya 01	saya ada, 01 ada,	2 dari 2	Saya 01	Benar
402	#CapresPilihanNetizen	Capres Pilihan Netizen	capres ada, pilihan ada, netizen ada,	3 dari 3	Capres Pilihan Netizen	Benar
403	#TrisaktiTurunLagi	Trisakti Turun Lagi	trisakti ada, turun ada, lagi ada,	3 dari 3	Trisakti Turun Lagi	Benar
404	#PetaniMakinSejahtera	Petani Makin Sejahtera	petani ada, makin ada, sejahtera ada,	3 dari 3	Petani Makin Sejahtera	Benar
405	#MingguTerakhir2019	Minggu Terakhir 2019	minggu ada, terakhir ada, 2019 ada,	3 dari 3	Minggu Terakhir 2019	Benar

406	#JiwasrayaSkandalPilpre	Jiwasraya Skandal Pilpres	jiwasraya ada, skandal ada, pilpres ada,	3 dari 3	Jiwasraya Skandal Pilpres	Benar
407	#KemajuanCepatDiPapua	Kemajuan Cepat Di Papua	kemajuan ada, cepat ada, di ada, papua ada,	4 dari 4	Kemajuan Cepat Di Papua	Benar
408	#IyainAjaBiarGaBanyakBacot	Iyain Aja Biar Ga Banyak Bacot	iyain ada, aja ada, biar ada, ga tidak ada, banyak ada, bacot ada,	5 dari 6	Iyain Aja Biar Ga Banyak Bacot	Benar
409	#TwitterMengundangRadit	Twitter Mengundang Radit	twitter ada, mengundang ada, radit ada,	3 dari 3	Twitter Mengundang Radit	Benar
410	#JanganBuatKamiBodoh	Jangan Buat Kami Bodoh	jangan ada, buat ada, kami ada, bodoh ada,	4 dari 4	Jangan Buat Kami Bodoh	Benar
411	#BenderaTauhidMilikIslam	Bendera Tauhid Milik Islam	bendera ada, tauhid ada, milik ada, islam ada,	4 dari 4	Bendera Tauhid Milik Islam	Benar
412	#HidupmuTerlaluBanyakDrama	Hidupmu Terlalu Banyak Drama	hidupmu ada, terlalu ada, banyak ada, drama ada,	4 dari 4	Hidupmu Terlalu Banyak Drama	Benar
413	#PergiHilangDanLupakan	Pergi Hilang Dan Lupakan	pergi ada, hilang ada, dan ada, lupakan ada,	4 dari 4	Pergi Hilang Dan Lupakan	Benar
414	#2019GantiPresiden	2019 Ganti Presiden	2019 ada, ganti ada, presiden ada,	3 dari 3	2019 Ganti Presiden	Benar
415	#RakyatSorotKeputusanMK	Rakyat Sorot Keputusan MK	rakyat ada, sorot ada, keputusan ada, mk ada,	4 dari 4	Rakyat Sorot Keputusan MK	Benar
416	#2019PrabowoPresidenFPI	2019 Prabowo Presiden FPI	2019 ada, prabowo ada, presiden ada, fpi ada,	4 dari 4	2019 Prabowo Presiden FPI	Benar
417	#IslamBerjayaKembali	Islam Berjaya Kembali	islam ada, berjaya ada, kembali ada,	3 dari 3	Islam Berjaya Kembali	Benar
418	#AniesBeliBeritaMediaOnline	Anies Beli Berita Media Online	anies ada, beli ada, berita ada, media ada, online ada,	5 dari 5	Anies Beli Berita Media Online	Benar

419	#KamiBersamaKPAI	Kami Bersama KPAI	kami ada, bersama ada, kpai ada,	3 dari 3	Kami Bersama KPAI	Benar
420	#BubarkanKPAI	bubarkan KPAI	bubarkan ada, kpai ada,	2 dari 2	bubarkan KPAI	Benar
421	#VisiIndonesia	Visi Indonesia	visi ada, indonesia ada,	2 dari 2	Visi Indonesia	Benar
422	#PercumaMainTwitter	Percuma Main Twitter	percuma ada, main ada, twitter ada,	3 dari 3	Percuma Main Twitter	Benar
423	#TahunBaruIslam	Tahun Baru Islam	tahun ada, baru ada, islam ada,	3 dari 3	Tahun Baru Islam	Benar
424	#TahunBaruImlek	Tahun Baru Imlek	tahun ada, baru ada, imlek ada,	3 dari 3	Tahun Baru Imlek	Benar
425	#RezimBaperan	Rezim Baperan	rezim ada, baperan ada,	2 dari 2	Rezim Baperan	Benar
426	#RakyatPilihJokowi	Rakyat Pilih Jokowi	rakyat ada, pilih ada, jokowi ada,	3 dari 3	Rakyat Pilih Jokowi	Benar
427	#AkuGampang	Aku Gampang	aku ada, gampang ada,	2 dari 2	Aku Gampang	Benar
428	#SandiwaraAnakMami	Sandiwara Anak Mami	sandiwara ada, anak ada, mami ada,	3 dari 3	Sandiwara Anak Mami	Benar
429	#SenjataDewanJenderal	Senjata Dewan Jenderal	senjata ada, dewan ada, jenderal ada,	3 dari 3	Senjata Dewan Jenderal	Benar
430	#PindahIbuKotaUntukSiapa	Pindah Ibu Kota Untuk Siapa	pindah ada, ibu ada, kota ada, untuk ada, siapa ada,	5 dari 5	Pindah Ibu Kota Untuk Siapa	Benar
431	#TolakJaksaAgungDariPartai	Tolak Jaksa Agung Dari Partai	tolak ada, jaksa ada, agung ada, dari ada, partai ada,	5 dari 5	Tolak Jaksa Agung Dari Partai	Benar
432	#SarapanSeru	Sarapan Seru	sarapan ada, seru ada,	2 dari 2	Sarapan Seru	Benar
433	#PerkembanganNyataPapua	Perkembangan Nyata Papua	perkembangan ada, nyata ada, papua ada,	3 dari 3	Perkembangan Nyata Papua	Benar
434	#SantuyTerakhir2019	Santuy Terakhir 2019	santuy ada, terakhir ada, 2019 ada,	3 dari 3	Santuy Terakhir 2019	Benar

435	#SemangatSelaluBuatKa	Semangat Selalu Buat Kamu	semangat ada, selalu ada, buat ada, kamu ada,	4 dari 4	Semangat Selalu Buat Kamu	Benar
436	#TepukTepungTawar	Tepuk Tepung Tawar	tepek ada, tepung ada, tawar ada,	3 dari 3	Tepuk Tepung Tawar	Benar
437	#THRBelumTerlihat	THR Belum Terlihat	thr ada, belum ada, terlihat ada,	3 dari 3	THR Belum Terlihat	Benar
438	#DiskualifikasiJokowi	Diskualifikasi Jokowi	diskualifikasi ada, jokowi ada,	2 dari 2	Diskualifikasi Jokowi	Benar
439	#KitaButuhKhalifah	Kita Butuh Khalifah	kita ada, butuh ada, khalifah ada,	3 dari 3	Kita Butuh Khalifah	Benar
440	#JakartaBeragam	Jakarta Beragam	jakarta ada, beragam ada,	2 dari 2	Jakarta Beragam	Benar
441	#MuslimKecewaAnies	Muslim Kecewa Anies	muslim ada, kecewa ada, anies ada,	3 dari 3	Muslim Kecewa Anies	Benar
442	#TangkapAbuJanda	Tangkap Abu Janda	tangkap ada, abu ada, janda ada,	3 dari 3	Tangkap Abu Janda	Benar
443	#TangkapPermadiArya	Tangkap Permadi Arya	tangkap ada, permadi ada, arya ada,	3 dari 3	Tangkap Permadi Arya	Benar
444	#PapuaBukanMonyet	Papua Bukan Monyet	papua ada, bukan ada, monyet ada,	3 dari 3	Papua Bukan Monyet	Benar
445	#KitaIniSama	Kita Ini Sama	kita ada, ini ada, sama ada,	3 dari 3	Kita Ini Sama	Benar
446	#BersamaTanganiKarhutla	Bersama Tangani Karhutla	bersama ada, tangani ada, karhutla ada,	3 dari 3	Bersama Tangani Karhutla	Benar
447	#MamahTahuSendiri	Mamah Tahu Sendiri	mamah ada, tahu ada, sendiri ada,	3 dari 3	Mamah Tahu Sendiri	Benar
448	#KTPGue	KTP Gue	ktp ada, gue ada,	2 dari 2	KTP Gue	Benar
449	#NongkrongBarengPrabowo	nongkrong bareng prabowo	nongkrong ada, bareng ada, prabowo ada,	3 dari 3	nongkrong bareng prabowo	Benar
450	#SatukanDangdut	Satukan Dangdut	satukan ada, dangdut ada,	2 dari 2	Satukan Dangdut	Benar
451	#SandiwaraUno	Sandiwara Uno	sandiwara ada, uno ada,	2 dari 2	Sandiwara Uno	Benar

452	#NetizenBersamaDigeembok	Netizen Bersama Digeembok	netizen ada, bersama ada, digeembok ada,	3 dari 3	Netizen Bersama Digeembok	Benar
453	#ZonaIndonesia	Zona Indonesia	zona ada, indonesia ada,	2 dari 2	Zona Indonesia	Benar
454	#AyoDengarRadio	Ayo Dengar Radio	ayo ada, dengar ada, radio ada,	3 dari 3	Ayo Dengar Radio	Benar
455	#MawarDanMelatiEps8	Mawar Dan Melati Eps 8	mawar ada, dan ada, melati ada, eps ada, 8 ada,	5 dari 5	Mawar Dan Melati Eps 8	Benar
456	#RomanHarusBertahan	Roman Harus Bertahan	roman ada, harus ada, bertahan ada,	3 dari 3	Roman Harus Bertahan	Benar
457	#2KataUntukMantan	2 Kata Untuk Mantan	2 ada, kata ada, untuk ada, mantan ada,	4 dari 4	2 Kata Untuk Mantan	Benar
458	#AntriTiketTimnas	Antri Tiket Timnas	antri ada, tiket ada, timnas ada,	3 dari 3	Antri Tiket Timnas	Benar
459	#KamiTakTakut	Kami Tak Takut	kami ada, tak ada, takut ada,	3 dari 3	Kami Tak Takut	Benar
460	#PrabowoBentengNKRI	Prabowo Benteng NKRI	prabowo ada, benteng ada, nkri ada,	3 dari 3	Prabowo Benteng NKRI	Benar
461	#01CapresPendustaAbadi	01 Capres Pendusta Abadi	01 ada, capres ada, pendusta ada, abadi ada,	4 dari 4	01 Capres Pendusta Abadi	Benar
462	#02CapresAbadi	02 Capres Abadi	02 ada, capres ada, abadi ada,	3 dari 3	02 Capres Abadi	Benar
463	#TandaDiaNggaSuka	Tanda Dia Ngga Suka	tanda ada, dia ada, ngga ada, suka ada,	4 dari 4	Tanda Dia Ngga Suka	Benar
464	#JokowiNyataKerjanya	Jokowi Nyata Kerjanya	jokowi ada, nyata ada, kerjanya ada,	3 dari 3	Jokowi Nyata Kerjanya	Benar
465	#JokowiMemberikanBukti	Jokowi Memberikan Bukti	jokowi ada, memberikan ada, bukti ada,	3 dari 3	Jokowi Memberikan Bukti	Benar
466	#NyoblosTanpaRusuh	Nyoblos Tanpa Rusuh	nyoblos ada, tanpa ada, rusuh ada,	3 dari 3	Nyoblos Tanpa Rusuh	Benar

467	#MenjemputRejeki2019	Menjemput Rejeki 2019	menjemput ada, rejeki ada, 2019 ada,	3 dari 3	Menjemput Rejeki 2019	Benar
468	#PilkadaDKI	Pilkada DKI	pilkada ada, dki ada,	2 dari 2	Pilkada DKI	Benar
469	#HitungCepatKompas	Hitung Cepat Kompas	hitung ada, cepat ada, kompas ada,	3 dari 3	Hitung Cepat Kompas	Benar
470	#HariPertamaSekolah	Hari Pertama Sekolah	hari ada, pertama ada, sekolah ada,	3 dari 3	Hari Pertama Sekolah	Benar
471	#KedaulatanRakyat	Kedaulatan Rakyat	kedaulatan ada, rakyat ada,	2 dari 2	Kedaulatan Rakyat	Benar
472	#TidakAtasNamaSaya	Tidak Atas Nama Saya	tidak ada, atas ada, nama ada, saya ada,	4 dari 4	Tidak Atas Nama Saya	Benar
473	#AyoSemuaBergerak	Ayo Semua Bergerak	ayo ada, semua ada, bergerak ada,	3 dari 3	Ayo Semua Bergerak	Benar
474	#TolakSementaraTurisCh	Tolak Sementara Turis China	tolak ada, sementara ada, turis ada, china ada,	4 dari 4	Tolak Sementara Turis China	Benar
475	#KabinetIndonesiaMaju	Kabinet Indonesia Maju	kabinet ada, indonesia ada, maju ada,	3 dari 3	Kabinet Indonesia Maju	Benar
476	#CeritaBaruSCTV	Cerita Baru SCTV	cerita ada, baru ada, sctv ada,	3 dari 3	Cerita Baru SCTV	Benar
477	#DukungJokowiBersihkanBUMN	Dukung Jokowi Bersihkan BUMN	dukung ada, jokowi ada, bersihkan ada, bumn ada,	4 dari 4	Dukung Jokowi Bersihkan BUMN	Benar
478	#EluYangSalahEluYangGalak	Elu Yang Salah Elu Yang Galak	elu ada, yang ada, salah ada, elu ada, yang ada, galak ada,	6 dari 6	Elu Yang Salah Elu Yang Galak	Benar
479	#AntiNgaret	Anti Ngaret	anti ada, ngaret ada,	2 dari 2	Anti Ngaret	Benar
480	#SuaraBuaya	Suara Buaya	suara ada, buaya ada,	2 dari 2	Suara Buaya	Benar
481	#CenayangDapatBasil	Cenayang Dapat Basil	cenayang ada, dapat ada, basil ada,	3 dari 3	Cenayang Dapat Basil	Benar
482	#IndonesiaDaruratAsap	Indonesia Darurat Asap	indonesia ada, darurat ada, asap ada,	3 dari 3	Indonesia Darurat Asap	Benar

483	#JokowiPilihanTepat	Jokowi Pilihan Tepat	jokowi ada, pilihan ada, tepat ada,	3 dari 3	Jokowi Pilihan Tepat	Benar
484	#BongkarDalangKm50	Bongkar Dalang Km 50	bongkar ada, dalang ada, km ada, 50 ada,	4 dari 4	Bongkar Dalang Km 50	Benar
485	#KeadilanDikebiri	Keadilan Dikebiri	keadilan ada, dikebiri ada,	2 dari 2	Keadilan Dikebiri	Benar
486	#AyoIstiqomahBerdakwa	Ayo Istiqomah Berdakwah	ayo ada, istiqomah ada, berdakwah ada,	3 dari 3	Ayo Istiqomah Berdakwah	Benar
487	#BeraniBerubah	Berani Berubah	berani ada, berubah ada,	2 dari 2	Berani Berubah	Benar
488	#DayakGanteng	Dayak Ganteng	dayak ada, ganteng ada,	2 dari 2	Dayak Ganteng	Benar
489	#VaksinCegahKematian	Vaksin Cegah Kematian	vaksin ada, cegah ada, kematian ada,	3 dari 3	Vaksin Cegah Kematian	Benar
490	#PercumaLaporPolisi	Percuma Lapor Polisi	percuma ada, lapor ada, polisi ada,	3 dari 3	Percuma Lapor Polisi	Benar
491	#TolakKampusAsing	Tolak Kampus Asing	tolak ada, kampus ada, asing ada,	3 dari 3	Tolak Kampus Asing	Benar
492	#7TahunJokowiCukup	7 Tahun Jokowi Cukup	7 ada, tahun ada, jokowi ada, cukup ada,	4 dari 4	7 Tahun Jokowi Cukup	Benar
493	#RepublikPencitraan	Republik Pencitraan	republik ada, pencitraan ada,	2 dari 2	Republik Pencitraan	Benar
494	#AniesMendunia	Anies Mendunia	anies ada, mendunia ada,	2 dari 2	Anies Mendunia	Benar
495	#PresidenTerburukDalamSejarah	Presiden Terburuk Dalam Sejarah	presiden ada, terburuk ada, dalam ada, sejarah ada,	4 dari 4	Presiden Terburuk Dalam Sejarah	Benar
496	#PernyataPartaiKorupsi	Ternyata Partai Korupsi	ternyata ada, partai ada, korupsi ada,	3 dari 3	Ternyata Partai Korupsi	Benar
497	#PercumaadaPolisi	percuma ada polisi	percuma ada, ada ada, polisi ada,	3 dari 3	percuma ada polisi	Benar
498	#1hari1oknum	1 hari 1 oknum	1 ada, hari ada, 1 ada, oknum ada,	4 dari 4	1 hari 1 oknum	Benar

499	#UngkapKebeharanKasu	Ungkap Kebeharan Kasu Indah	ungkap ada, kebeharan ada, kasu ada, indah ada,	4 dari 4	Ungkap Kebeharan Kasu Indah	Benar
500	#NyanyiBareng	Nyanyi Bareng	Nyanyi ada, Bareng ada,	2 dari 2	Nyanyi Bareng	Benar

UIN SUSKA RIAU



#NyanyiBareng

Hak Cipta Dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

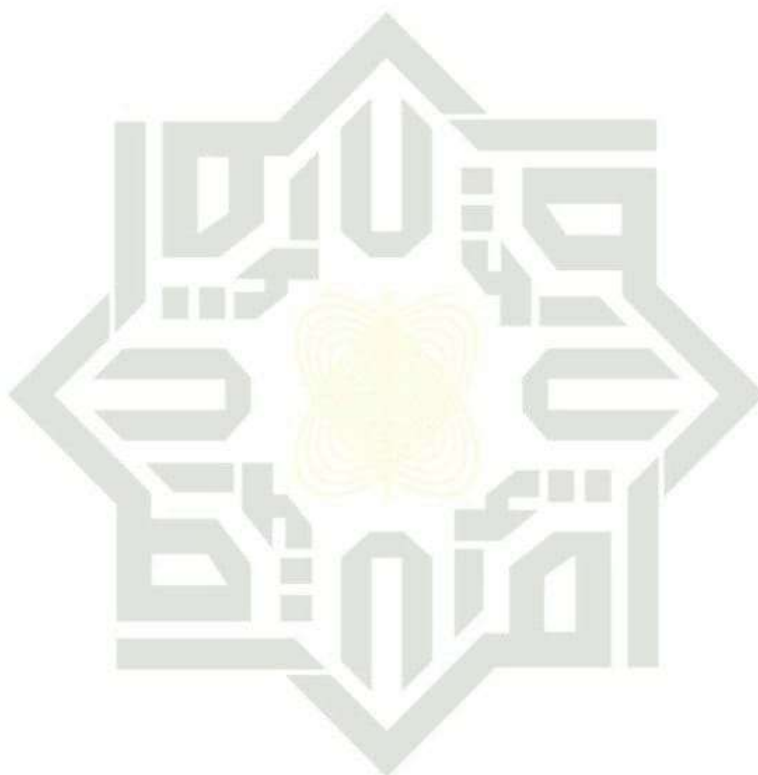
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

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

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

Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau



UIN SUSKA RIAU



2. Ditinjau mengenai...
a. ...
b. ...

2. Ditinjau mengenai...
a. ...
b. ...

an Syarif Kasim Riau
u tinjauan suatu masalah.

LAMPIRAN B

LEMBAR VALIDASI DATA IDENTIFIKASI WORD-BOUNDARY



LAMPIRAN C

DAFTAR RIWAYAT HIDUP



Hak Cipta dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
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.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© Himpunan Cipta m... UIN Suska Riau

Blas... of Sultan Syarif Kasim Riau

Informasi Personal

Nama : Ulfa Nada Diana
Tempat/Tanggal Lahir : Bangkinang / 28 Des 1996
Jenis Kelamin : Perempuan
Agama : Islam
Anak ke : 3 dari 3 bersaudara
Alamat : Jl. Jend. Sudirman Gg.Rahman
Bangkinang Kota
No HP : 082288084022
Email : ulfa.nada.diana@students.uin-suska.ac.id



Riwayat Pendidikan

Tahun 2001-2002 : TK Pembina Bangkinang Kota
Tahun 2002-2008 : SDN 002 Langgini, Bangkinang Kota
Tahun 2008-2012 : Pondok Pesantren Daarun Nahdhah Thawalib Bangkinang (Kelas 1-4)
Tahun 2012-2015 : SMAN 1 Bangkinang Kota
Tahun 2015-2022 : S1 Jurusan Teknik Informatika, Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau