



**IMPLEMENTASI BIDIRECTIONAL ENCODER  
REPRESENTATIONS FROM TRANSFORMERS (BERT)  
UNTUK MENDETEKSI HATESPEECH DAN ABUSIVE  
LANGUAGE PADA TWITTER BAHASA INDONESIA**

**JURUSAN TEKNIK INFORMATIKA  
UIN SUSKA RIAU  
TUGAS AKHIR**

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

Oleh

**RESKI SAPUTRA**

**NIM. 11651101881**



UIN SUSKA RIAU

UIN SUSKA RIAU

**FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU  
PEKANBARU**

**2022**

**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.



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.

## LEMBAR PERSETUJUAN

**IMPLEMENTASI BIDIRECTIONAL ENCODER  
REPRESENTATIONS FROM TRANSFORMERS (BERT)  
UNTUK MENDETEKSI HATESPEECH DAN ABUSIVE  
LANGUAGE PADA TWITTER BAHASA INDONESIA  
JURUSAN TEKNIK INFORMATIKA  
UIN SUSKA RIAU**

### TUGAS AKHIR

Oleh

**RESKI SAPUTRA**

**NIM. 11651101881**

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

Pembimbing I,

**SURYA AGUSTIAN, S.T, M.Kom**

**NIP. 19760830 201101 1 003**



**LEMBAR PENGESAHAN**

**IMPLEMENTASI BIDIRECTIONAL ENCODER  
REPRESENTATIONS FROM TRANSFORMERS (BERT)  
UNTUK MENDETEKSI HATESPEECH DAN ABUSIVE  
LANGUAGE PADA TWITTER BAHASA INDONESIA  
JURUSAN TEKNIK INFORMATIKA  
UIN SUSKA RIAU**

Oleh

**RESKI SAPUTRA**

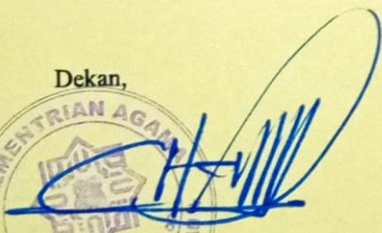
**NIM. 11651101881**

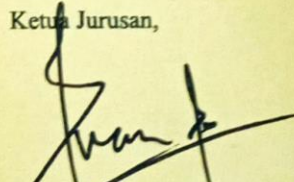
Telah dipertahankan di depan sidang dewan penguji  
sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik  
pada Universitas Islam Negeri Sultan Syarif Kasim Riau

Pekanbaru, 7 Januari 2022

Mengesahkan,

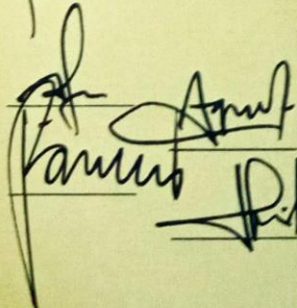
Ketua Jurusan,

Dekan,  
  
**Dr. HARTONO, M.Pd**  
**NIP. 19640301 199203 1 003**

  
**IWAN ISKANDAR, M.T**  
**NIP. 19821216 201503 1 003**

**DEWAN PENGUJI**

Ketua : Muhammad Affandes, M.T  
Pembimbing I : Surya Agustian, S.T, M.Kom.  
Penguji I : Benny Sukma Negara, M.T  
Penguji II : Suwanto Sanjaya, ST, M.Kom.



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.

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

### SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini:

Nama : RESKI SAPUTRA  
 NIM : 11651101801  
 Tempat/Tgl. Lahir : Pekanbaru, 23 Januari 1997  
 Fakultas/Pascasarjana : Sains dan Teknologi  
 Prodi : Teknik Informatika  
 Judul Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya\*:

Implementasi Bidirectional Encoder Representations From Transformers (BERT) untuk Mendeteksi Hate Speech dan Abusive Language pada Twitter Bahasa Indonesia

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, 12 Januari 2022  
 Yang membuat pernyataan

  
  
 Reski Saputra  
 NIM : 11651101801

\*pilih salah satu sesuai jenis karya tulis

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.



## 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 dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan 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.

### 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.



## 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 terdapat dalam naskah ini dan disebutkan didalam daftar pustaka.

Pekanbaru, 7 Januari 2022

Yang membuat pernyataan,

**RESKI SAPUTRA**

**NIM. 11651101881**

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.



## LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Alhamdulillah Rabbil ‘Alamin, puji syukur yang teramat dalam kepada Allah SWT, yang telah memberikan salah satu anugerah terindah yang tidak akan pernah terlupakan seumur hidup.

Kepada kedua Orang Tua tercinta, terima kasih selalu berjuang, memberikan kasih sayang dan do’a yang tak ternilai harganya, yang akan selalu dibutuhkan kapanpun dan dimanapun. Dari lubuk hati yang paling dalam, ku persembahkan gelar sarjan ini.

Kepada keluarga tersayang, terima kasih telah menjadi pendengar yang baik, memberi motivasi serta semangat yang sangat berarti bagi penulis hingga penulis bisa sampai ke titik ini.

Kepada teman-teman seperjuangan, terima kasih karena selalu ada dan berjuang bersama. Mari terus berjuang bersama-sama, karena bersama kita bisa.

### 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.



## ABSTRAK

Ujaran kebencian dan bahasa kasar di media sosial merupakan salah satu penyebab terjadinya konflik di masyarakat. Terdapatnya ujaran kebencian dan bahasa kasar pada media sosial dikarenakan pada media sosial pengguna dapat dengan bebas untuk menyampaikan pendapatnya, oleh karena itu konten ujaran kebencian dan bahasa kasar pada media sosial perlu dideteksi dan dibatasi. Pada penelitian [1] telah dilakukan penelitian dalam mendeteksi *hatespeech* dan *abusive* beserta target, kategori dan level *hatespeech* menggunakan berbagai feature extraction, classifier dan transformasi data. Pada penelitian tersebut *word unigram*, *Random Forest Decision Tree* dan *label power-set* merupakan kombinasi terbaik dengan akurasi 66.12%. Pada penelitian tersebut belum didapatkan hasil yang optimal dalam mendeteksi *hatespeech* dan *abusive* beserta target, kategori dan level *hatespeech*. Beberapa tahun belakangan ini *neural network* yang dikombinasikan dengan *pretained language model* seperti *Bidirectional encoder from transformers* (BERT) mendapatkan akurasi yang baik dalam berbagai tugas *natural language processing*. Pada penelitian ini dilakukan penelitian dengan membuat model *neural network* dengan BERT untuk mengklasifikasi *hatespeech* dan *abusive language* beserta target, kategori dan level. Hasil dari penelitian ini didapatkan bahwa model *neural network* dengan BERT mendapatkan hasil yang lebih baik dari penelitian sebelumnya yaitu dengan akurasi 72.28%

Kata kunci: Ujaran kebencian, Bahasa kasar, Bidirectional encoder from transformer, Klasifikasi, twitter.

### 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.





## ABSTRACT

Hate speech and abusive language on social media is one of the causes of conflict in society. The presence of hate speech and abusive language on social media is because on social media users can freely express their opinions, therefore the content of hate speech and abusive language on social media needs to be detected and limited. In research [1] research has been carried out in detecting hate speech and abusive along with targets, categories and levels of hate speech using various feature extraction, classifier and data transformation. In this study word unigrams, Random Forest Decision Trees and power-set labels were the best combination with an accuracy of 66.12%. In this study, optimal results have not been obtained in detecting hate speech and abusive along with the target, category and level of hate speech. In recent years, neural networks combined with preserved language models such as Bidirectional encoder from transformers (BERT) have achieved good accuracy in various natural language processing tasks. In this study, research was conducted by making a neural network model with BERT to classify hate speech and abusive language along with targets, categories and levels. The results of this study showed that the neural network model with BERT got better results than previous studies with an accuracy of 72.28%.

Keywords: Hate speech, Abusive language, Bidirectional encoder from transformer, classification, twitter.

### 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. 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.

## KATA PENGANTAR

*Assalammu 'alaikum wa rohmatullohi wa barokatuh.*

*Athamdulillahi robbil'alamin*, tak henti-hentinya kami ucapkan kehadiran Allah *Subhanahu wa ta'ala*, yang dengan rahmat dan hidayah-Nya kami mampu menyelesaikan Tugas Akhir ini dengan baik. Tidak lupa bershalawat kepada Nabi dan Rasul-Nya, Nabi Muhammad *Sholallohu 'alaihi wa salam*, yang telah membimbing kita sebagai umatnya menuju jalan kebaikan.

Tugas Akhir ini disusun sebagai salah satu syarat untuk mendapatkan gelar sarjana pada jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau. Banyak sekali pihak yang telah membantu kami 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 kami hanya dapat mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Hairunas, 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, M.T, selaku Ketua Jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Ibu Lola Oktavia, S.S.T, M.T.I., selaku Pembimbing Akademis penulis selama menjalani perkuliahan di Jurusan Teknik Informatika, yang telah banyak memberikan arahan serta masukan mengenai perkuliahan bagi penulis.
5. Bapak Surya Agustian, S.T, M.Kom, selaku dosen pembimbing Tugas Akhir penulis, yang telah sangat banyak berbagi waktu, ilmu dan wawasan yang dimiliki kepada penulis sehingga penelitian dan Laporan Tugas Akhir ini selesai.



**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.

6. Bapak Benny Sukma Negara, S.T, M.T, selaku dosen penguji I yang telah meluangkan waktunya dan banyak memberikan saran, dan masukan yang membangun bagi penulis.
7. Bapak Suwanto Sanjaya, ST, M.Kom, selaku dosen penguji II yang telah meluangkan waktunya dan banyak memberikan saran, dan masukan yang membangun bagi penulis.
8. Seluruh Bapak/Ibu dosen Teknik Informatika yang telah sabar memberikan tunjuk ajar serta ilmu yang bermanfaat kepada penulis selama masa perkuliahan.
9. Keluarga penulis yang tidak pernah letih memberikan semangat, nasehat, dan doa yang tulus.
10. Teman-teman seperjuangan angkatan 2016 yang berjuang dan saling mendukung.
11. Seluruh pihak yang belum kami cantumkan, terima kasih atas dukungannya, baik material maupun spiritual.

Kami 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 kami berharap semoga laporan ini dapat memberikan sesuatu yang bermanfaat bagi siapa saja yang membacanya.

*Wassalamu'alaikum wa rohmatullohi wa barokatuh.*

Pekanbaru, 7 Januari 2022

Penulis

UIN SUSKA RIAU



## DAFTAR ISI

LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN .....	iii
LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL.....	iv
LEMBAR PERNYATAAN .....	v
LEMBAR PERSEMBAHAN .....	vi
ABSTRAK .....	vii
ABSTRACT .....	viii
KATA PENGANTAR .....	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR .....	xiv
DAFTAR TABEL.....	xvi
BAB 1 PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	5
1.3 Batasan Masalah.....	6
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	6
BAB 2 KAJIAN PUSTAKA.....	7
2.1 Kajian Metode .....	7
2.1.1 Twitter.....	7
2.1.2 <i>Hatespeech</i> .....	7
2.1.3 <i>Abusive Language</i> .....	8
2.1.4 <i>Artificial Inteligence</i> .....	8

### 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. 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.

2.1.5	<i>Machine Learning</i> .....	9
2.1.6	<i>Artificial Neural Network (ANN)</i> .....	10
2.1.7	<i>Deep Learning</i> .....	13
2.1.8	<i>Transformer</i> .....	13
2.1.9	<i>Bidirectional Encoder Representations from Transformers (BERT)</i> 16	
2.1.10	<i>Transfer Learning</i> .....	19
2.1.11	<i>Confusion Matrix</i> .....	19
2.2	Penelitian Terkait .....	20
<b>BAB 3 METODOLOGI PENELITIAN</b> .....		25
3.1	Identifikasi Masalah .....	25
3.2	Studi Literatur.....	26
3.3	Pengumpulan Data .....	26
3.4	Analisa.....	29
3.4.1	Analisa Dataset.....	30
3.4.2	<i>Teks Preprocessing</i> .....	30
3.4.3	Analisa BERT .....	31
3.4.4	<i>Fine tuning BERT</i> .....	32
3.5	Implementasi .....	33
3.6	Pengujian .....	33
3.7	Kesimpulan dan Saran.....	33
<b>BAB 4 PEMBAHASAN</b> .....		34
4.1	Analisa.....	34
4.1.1	Analisa Dataset.....	34
4.1.2	Teks Preprocessing.....	40



**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.

4.1.3	Analisa BERT .....	45
4.1.4	Fine tuning BERT .....	45
4.2	Perancangan.....	57
4.3	Skenario Eksperimen.....	57
4.4	Implementasi .....	60
4.4.1	Batasan Implementasi .....	60
4.4.2	Lingkungan Implementasi.....	60
4.4.3	Implementasi model klasifikasi menggunakan BERT(bidirectional encoder from transformers).....	61
4.5	Pengujian .....	74
4.5.1	Pengujian Hyperparameter.....	74
4.5.2	Pengujian Teks Preprocessing .....	78
4.5.3	Confusion Matrix .....	79
4.5.4	Hasil Pengujian .....	85
BAB 5 PENUTUP .....		87
5.1	Kesimpulan.....	87
5.2	Saran .....	87
DAFTAR PUSTAKA .....		89
LAMPIRAN A .....		93
LAMPIRAN B .....		96
DAFTAR RIWAYAT HIDUP.....		99



## DAFTAR GAMBAR

Gambar 1 Neuron pada Otak Manusia.....	11
Gambar 2 Struktur Artificial neural network.....	12
Gambar 3 Arsitektur Encoder Decoder pada Transformer.....	14
Gambar 4 The Transformer - Model Architecture.....	15
Gambar 5 Scaled Dot-Product Attention dan Multi-Head Attention.....	15
Gambar 6 Arsitektur BERT.....	17
Gambar 7 Tahapan Metodologi Penelitian.....	25
Gambar 8 Pembagian Dataset Hatespeech dan Abusive Language.....	27
Gambar 9 Tahapan Analisa Model BERT.....	30
Gambar 10 Tahapan <i>fine-tuning</i> BERT.....	32
Gambar 11 Jumlah label pada dataset.....	35
Gambar 12 Jumlah label pada dataset ( <i>undersampling</i> ).....	37
Gambar 13 Model Klasifikasi BERT.....	46
Gambar 14 Feedforward Neural Network.....	54
Gambar 15 Output Proses Klasifikasi.....	56
Gambar 16 Flowchart Perancangan Model BERT.....	57
Gambar 17 Import Library Python.....	63
Gambar 18 Import Dataset.....	63
Gambar 19 Cek Data Kosong.....	64
Gambar 20 Kode program grafik jumlah data setiap label.....	65
Gambar 21 Grafik jumlah data setiap label.....	65
Gambar 22 Pembagian data latih dan data uji.....	66
Gambar 23 Kode Program Case Folding.....	66
Gambar 24 Kode Program Cleaning.....	67
Gambar 25 Kode Program Normalisasi.....	67
Gambar 26 Kode Program Menghapus Stopword.....	67
Gambar 27 Kode Program Penerapan Teks Preprocessing pada dataset.....	68
Gambar 28 Kode program tokenisasi data.....	69

### Hak Cipta Dilindungi Undang-Undang

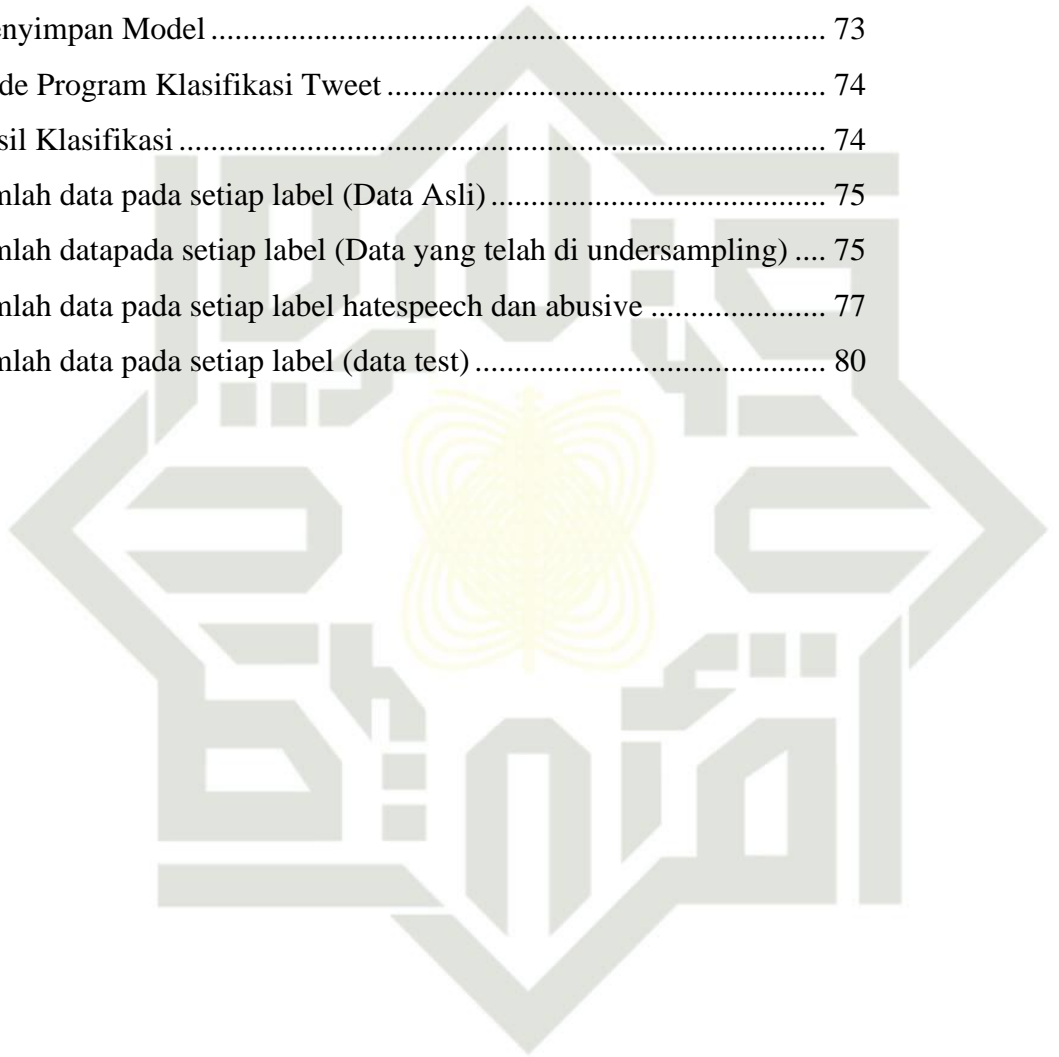
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan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

Gambar 29 Pengecekan Panjang maksimal token pada dataset.....	69
Gambar 30 Kode Program Tokenisasi Seluruh Data .....	70
Gambar 31 Kode Program Persiapan Data .....	71
Gambar 32 Kode Program Model Klasifikasi BERT .....	72
Gambar 33 Proses training BERT .....	73
Gambar 34 Menyimpan Model .....	73
Gambar 35 Kode Program Klasifikasi Tweet .....	74
Gambar 36 Hasil Klasifikasi .....	74
Gambar 37 Jumlah data pada setiap label (Data Asli) .....	75
Gambar 38 Jumlah data pada setiap label (Data yang telah di undersampling) ....	75
Gambar 39 Jumlah data pada setiap label hatespeech dan abusive .....	77
Gambar 40 Jumlah data pada setiap label (data test) .....	80







## DAFTAR TABEL

Tabel 1 <i>Confusion Matrix</i> .....	20
Tabel 2 Penelitian Terkait .....	20
Tabel 3 Jumlah Seluruh Label .....	34
Tabel 4 Jumlah Label pada Target, Kategori dan Level Hatespeech.....	36
Tabel 5 Dataset Awal yang akan digunakan .....	37
Tabel 6 Hasil Case folding .....	40
Tabel 7 Hasil Cleaning .....	41
Tabel 8 Hasil Normalization .....	42
Tabel 9 Hasil Stopword Removal .....	43
Tabel 10 Hasil Wordpiecer Tokenizer .....	44
Tabel 11 Data Training .....	46
Tabel 12 Inisialisasi Parameter .....	53
Tabel 13 Skenario Eksperimen Hyperparameter .....	58
Tabel 14 Skenario Eksperimen Teks Preprocessing .....	59
Tabel 15 Library Python .....	62
Tabel 16 Hasil Pengujian Kombinasi Hyperparameter.....	76
Tabel 17 Hasil Pengujian hyperparameter terbaik pada dataset hatespeech dan abusive.....	78
Tabel 18 Hasil Pengujian Eksperimen Teks Preprocessing.....	78
Tabel 19 Confusion matrix label Hatespeech .....	80
Tabel 20 Confusion matrix label Abusive .....	81
Tabel 21 Confusion matrix label HS_Individual .....	81
Tabel 22 Confusion matrix label HS_Group .....	81
Tabel 23 Confusion matrix label HS_Religion.....	82
Tabel 24 Confusion matrix label HS_Race.....	82
Tabel 25 Confusion matrix label HS_Physical .....	83
Tabel 26 Confusion matrix label HS_Gender .....	83
Tabel 27 Confusion matrix label HS_Other .....	83

- 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. 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.

Tabel 28 Confusion matrix label HS_Weak .....	84
Tabel 29 Confusion matrix label HS_Moderate .....	84
Tabel 30 Confusion matrix label HS_Strong .....	85



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.

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan teknologi informasi terutama dalam bidang media dan komunikasi membuat orang-orang dengan mudah untuk mengakses dan membagikan informasi [2]. Salah satu teknologi yang terus berkembang adalah media sosial. Media sosial menjadi platform yang banyak digunakan dan terus bertumbuh setiap tahunnya. Menurut *we are sosial* pengguna media sosial di Indonesia pada tahun 2020 berjumlah 160 juta dan bertumbuh sebesar 8.1% dari tahun sebelumnya [3]. Media sosial membuat informasi dengan cepat tersebar diseluruh dunia, menjadikannya tidak hanya digunakan sebagai sarana pertemanan dan berbagi informasi tetapi juga digunakan sebagai sarana untuk perdagangan, sosialisasi kebijakan pemerintah, kampanye politik dan lainnya. singkatnya media sosial sekarang dapat digunakan untuk tujuan apa pun [4].

Media sosial membuat orang-orang dengan mudah untuk membagikan pendapat dan berekspresi dengan bebas, hal ini memungkinkan terjadinya penyebaran informasi yang mengandung ujaran kebencian [5]. Ujaran kebencian adalah aksi komunikasi yang dilakukan oleh orang ataupun kelompok dalam wujud provokasi, hasutan, maupun hinaan dengan tujuan untuk menjatuhkan orang ataupun kelompok lain dalam berbagai aspek seperti warna kulit, ras, etnis, gender, cacatan, orientasi seksual, kewarganegaraan, agama dan lain-lain [6]. Dalam media sosial ujaran kebencian seringkali disertai dengan bahasa kasar. Bahasa kasar adalah ungkapan yang mengandung kata-kata kasar atau kotor yang disampaikan baik secara lisan ataupun tulisan. Ujaran kebencian yang disertai bahasa kasar sering mempercepat terjadinya konflik sosial karena bahasa kasar dapat menyinggung perasaan dan memicu emosi [7]. Ujaran kebencian dan bahasa kasar dapat mendorong kekerasan atau mempromosikan permusuhan antar kelompok yang berbeda atas dasar agama, ras, tempat tinggal, bahasa dan lainnya. Hal ini

#### 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.

menjadikan ujaran kebencian dan bahasa kasar di media sosial sebagai penyebab terjadinya konflik sosial, pertentangan dan perpecahan di masyarakat [8]. Dengan terus berkembangnya media sosial membuat konten yang mengandung ujaran kebencian dan bahasa kasar menjadi cepat tersebar dan dapat dilihat oleh berbagai pengguna di seluruh dunia [4]. Oleh karena itu konten yang mengandung ujaran kebencian dan bahasa kasar pada media sosial perlu untuk dibatasi dan dideteksi.

Beberapa tahun belakangan ini telah banyak dilakukan penelitian tentang mendeteksi ujaran kebencian dan bahasa kasar di berbagai jenis bahasa dan media sosial. Pada penelitian [1], telah dilakukan penelitian dalam mendeteksi *hatespeech* dan *abusive language* beserta target, kategori dan level *hatespeech*. Pada penelitian tersebut pendeteksian target, kategori dan level bertujuan untuk membantu pihak berwajib dalam menentukan kasus ujaran kebencian yang perlu ditangani terlebih dahulu. Pada penelitian tersebut *dataset* yang digunakan merupakan *dataset hatespeech* dan *abusive language* yang di dapat dari twitter. Pada *dataset* tersebut terdapat beberapa label yaitu *hatespeech*, *abusive*, target *hatespeech*, kategori *hatespeech* dan level *hatespeech*. Setiap data yang ada pada *dataset* tersebut memiliki lebih dari satu label atau disebut dengan *multi-label*. Pada *dataset* tersebut suatu *tweet* bisa saja termasuk *hatespeech* tetapi *tidak abusive*, bisa juga tidak *hatespeech* tetapi *abusive* dan bisa juga termasuk *hatespeech* dan *abusive*. Selain itu pada *dataset* tersebut setiap *tweet* dengan label *hatespeech* memiliki target, kategori dan level.

Pada penelitian [1] peneliti melakukan beberapa eksperimen dalam mengklasifikasi *hatespeech* dan *abusive language* dengan menggunakan beberapa algoritma machine learning yaitu Naive Bayes, Support Vector Machine dan Random Forest Decision Tree yang di kombinasikan dengan beberapa feature extraction seperti *term frequency*(TF), *orthography* and *lexicon features*. *term frequency* yang digunakan adalah *word n-gram* (unigram, bigram dan trigram). Untuk *orthography feature* menggunakan jumlah dari tanda baca, tanda tanya, huruf kecil dan besar. Untuk *lexicon feature* menggunakan *sentiment lexicon* (sentimen negatif dan Positif). Ketiga algoritma tersebut merupakan algoritma klasifikasi

#### 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.

untuk *single* label, untuk dapat digunakan dalam menyelesaikan tugas *multi-label classification* digunakan metode transformasi data *Binary Relevance* (BR), *Label Power-set* (LP), and *Classifier Chains* (CC).

Dari beberapa *experiment* yang telah dilakukan pada penelitian [1] metode Random forest decision tree yang dikombinasikan dengan feature extraction *term frequency* (*word n-gram*) dan *Label Power-set* (LP) merupakan kombinasi *feature extraction*, metode dan transformasi data terbaik yang mendapatkan akurasi 77.36% dalam melakukan *multi-label clasification hatespeech dan abusive language*. Dari semua *experiment* yang telah dilakukan pada penelitian tersebut belum mendapatkan akurasi yang baik dalam melakukan *multi-label classification hatespeech dan abusive language* dengan identifikasi target, kategori dan *level hatespeech* dengan akurasi 66.12%. Pada penelitian tersebut peneliti menjelaskan terdapat banyak false negatif pada pengujian yang mungkin disebabkan oleh jumlah dataset yang tidak seimbang (*unbalanced dataset*).

Pada penelitian [1] peneliti memberikan saran untuk menggunakan semantic feature pada proses *feature extraction* seperti word embedding. Karena pada beberapa *experiment* pada bahasa indonesia [9] dengan menambahkan word embedding dapat meningkatkan kinerja klasifikasi karena pada *word embedding* dapat mengenali arti dari suatu kata yang tidak dapat dikenali oleh *feature* seperti *term frequency*, *orthography* dan *lexicon feature*. Word embedding merupakan teknik feature extraction yang merepresentasikan setiap kata menjadi suatu vector yang dan pada setiap kata yang memiliki arti yang hampir sama akan memiliki jarak vector yang dekat. Pada feature extraction biasa seperti *term frequency* (TF), *feature* dari suatu kata didapatkan berdasarkan jumlah kata yang muncul terhadap labelnya, sehingga jika suatu kata banyak muncul di suatu label yang sama maka kata tersebut memiliki hubungan yang kuat terhadap label tersebut [10]. Misalkan kata “kucing” sering muncul pada kalimat yang memiliki label *abusive* maka kata “kucing” memiliki nilai probabilitas yang besar terhadap label *abusive*, Oleh karena itu feature extraction biasa seperti *term frequency* tidak melihat suatu kata berdasarkan semantik atau arti kata tetapi hanya dengan jumlah kemunculan kata.

#### 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.

Beberapa tahun belakangan ini terdapat teknik *word embedding* yang dapat mengenali konteks pada kata, teknik ini disebut dengan *pretraining language model*. *Pretraining language model* mendapatkan hasil yang efektif untuk meningkatkan berbagai tugas *natural language processing* [11]. *Pretraining language model* merupakan model yang di latih dari data teks yang tidak berlabel dalam jumlah banyak untuk memahami makna dan relasi dari suatu kata [12]. Terdapat beberapa *pretraining language model* seperti *Universal language Model fine-tuning (ULMFiT)*, *embedding from language Models (ELMO)*, *OpenAI Generative Pre-trained Transformer (GPT)*, dan *Google BERT*. Dari beberapa *pretraining language model* tersebut BERT mendapatkan performa terbaik [11]. *Pretraining language model* yang telah dibuat dapat digunakan untuk berbagai tugas *natural language processing* seperti *classification*, *question answering*, *translation* dan lainnya yang dikenal dengan proses “*transfer learning*”. Dengan proses *transfer learning*, model yang telah dibuat dapat digunakan kembali dan dimodifikasi, sehingga proses pembuatan model machine learning NLP dapat menghemat waktu, energi dan *resources* karena proses ini tidak perlu melakukan *training language model* dari awal. Saat ini *pretrain BERT* tersedia dalam beberapa bahasa. Awalnya model BERT hanya dilatih menggunakan *dataset* teks yang besar dalam bahasa inggris. Pada penelitian [13] dilakukan proses *pretraining BERT* menggunakan *dataset* berbahasa indonesia. *Pretrain BERT* ini dilatih menggunakan teks data berukuran lebih kurang 23 GB, *dataset* didapat dari sumber seperti berita, sosial media, wikipedia, artikel, subtitles dari rekaman video dan *parallel dataset* yang disebut dengan *indoBERT*.

Beberapa penelitian sebelumnya yang digunakan sebagai referensi tentang klasifikasi *hatespeech* dan *abusive language*. Penelitian [12] yang berjudul *A BERT-Based Transfer Learning Approach for Hate Speech Detection in Online Social Media*. Penelitian ini menggunakan *pre-trained language model BERT* yang dibentuk dari data wikipedia dan *bookcorpus* berbahasa inggris. Pada penelitian ini dilakukan beberapa *fine-tuning strategies* model BERT seperti *BERT based*, *insert non linear layer*, *insert Bi-LSTM layer* dan *insert CNN layer*. Pada penelitian ini model BERT+CNN mendapatkan akurasi terbaik yaitu 92%.

#### 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.

Penelitian selanjutnya [14] yang berjudul Analisis Sentimen Review Film Berbahasa Inggris Dengan Pendekatan *Bidirectional Encoder Representations from Transformers*. Penelitian ini menggunakan *movie review polarity dataset v2.0* berbahasa inggris yang berjumlah 2000 data. 1000 data berlabel positif dan 1000 berlabel negatif. Sebelum menggunakan metode BERT, penulis membuat model klasifikasi menggunakan algoritma naïve bayes sebagai *baseline* dengan data *train*, data *test* dan *preprocessing* yang sama dengan yang digunakan pada saat membuat model metode BERT dan memperoleh hasil akurasi sebesar 48%. selanjutnya peneliti melakukan *fine-tuning* model BERT dengan menambah Bi-LSTM layer dan mendapatkan akurasi 78%.

Penelitian selanjutnya [15] yang berjudul *Hate Speech and abusive language classification using fastext*. Penelitian ini membangun beberapa model *hate speech classification* menggunakan *continous bag of words* dan fastteks. Pada penelitian ini model fastteks mendapatkan hasil yang lebih baik pada *binary clasification* tetapi tidak menampilkan hasil yang signifikan dari penelitian sebelumnya. Pada penelitian ini penulis menyarankan untuk membuat model BERT untuk mengklasifikasi *hate speech* dan *abusive language*.

Berdasarkan permasalahan diatas penulis melakukan penelitian tugas akhir dengan menerapkan *Bidirectional Encoder Representations from Transformers* untuk mengklasifikasi *hatespeech* dan *abusive language* pada Twitter berbahasa Indonesia. Tujuan dari penelitian ini adalah untuk meningkatkan akurasi dari penelitian sebelumnya dan melakukan beberapa experiment untuk mencari model BERT terbaik dan mengetahui tingkat akurasinya.

## 1.2 Rumusan Masalah

Berdasarkan permasalahan pada latar belakang penulis mengambil rumusan masalah, yaitu bagaimana menerapkan *Bidirectional Encoder from transformers* (BERT) untuk mengklasifikasi multi-label *hatespeech* dan *abusive language* pada Twitter bahasa indonesia serta mencari model terbaik dan melihat tingkat akurasi yang didapatkan.

#### 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.

### 1.3 Batasan Masalah

Berdasarkan rumusan masalah, maka dibutuhkan sebuah batasan masalah untuk membatasi ruang lingkup pembahasan yang akan dilakukan. Ruang lingkup yang akan dibahas yaitu:

1. *Dataset* yang diproses sebanyak 13.127 *tweet* [1]
2. Label yang diproses yaitu *hatespeech*, *abusive* dan *level hatespeech* (*weak*, *moderate*, *strong*). target (individual, group), categories (religion/creed, race/ethnicity, physical/disability, gender/sexual, other invective/slander) dan level(weak, moderate, strong).
3. Metode klasifikasi yang digunakan adalah *feedforward neural network* dengan *Bidirectional Encoder from transformers (BERT)*
4. Model BERT yang digunakan adalah IndoBERT [13]

### 1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini :

1. Implementasi *Bidirectional Encoder from transformers (BERT)* untuk mengklasifikasi *hatespeech* dan *abusive language* pada Twitter bahasa Indonesia
2. Mencari model BERT terbaik dan menghitung akurasi dari model *Bidirectional Encoder from transformers (BERT)* dalam mengklasifikasi *hatespeech* dan *abusive language* pada Twitter bahasa indonesia.

### 1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini antara lain:

1. Mengetahui penerapan, kinerja dan tingkat akurasi dari model *Bidirectional Encoder from transformers (BERT)* dalam mengklasifikasi *hatespeech* dan *abusive language* pada Twitter bahasa indonesia.
2. Menjadi sumber referensi dalam melakukan penelitian lanjutan khususnya pada klasifikasi *hatespeech* dan *abusive language* dan penerapan model *Bidirectional Encoder from transformers (BERT)*.



**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.

## BAB 2

### KAJIAN PUSTAKA

#### 2.1 Kajian Metode

##### 2.1.1 Twitter

Menurut halaman resmi Twitter.com, Twitter adalah media sosial yang menghubungkan orang-orang di seluruh dunia, twitter memungkinkan pengguna untuk berbicara satu sama lain dalam pesan pendek sepanjang 280 karakter yang disebut tweet. Twitter sendiri memungkinkan pengguna untuk mencari berita terbaru tentang topik yang diminati. Twitter mempunyai fitur utama yaitu dapat menuliskan status atau cuitan dan mengirim pesan ke pengguna lain. Fitur lain dari media sosial Twitter meliputi:

1. Hashtag (#) adalah sebuah penanda yang dibuat oleh pengguna Twitter untuk mengelompokan tweet pada sebuah topik tertentu. Hashtag dapat memudahkan pengguna untuk menemukan tweet tentang topik yang ingin mereka cari..
2. Trending topik adalah fitur yang menampilkan topik-topik atau hashtag yang sedang populer tweet atau banyak dibicarakan oleh pengguna. Adanya trending topik memungkinkan pengguna mengetahui apa yang sedang viral di masyarakat.
3. Retweet adalah fitur yang digunakan pengguna untuk me-retweet atau membagikan tweet pengguna lain ke halaman beranda priadinya.
4. Reply adalah fitur untuk membalas tweet dari pemilik akun Twitter lain. Untuk membalas Tweet seseorang, dapat dilakukan dengan menempatkan simbol @ dan diikuti dengan nama pengguna.

##### 2.1.2 Hatespeech

Pengertian sederhana tentang ujaran kebencian adalah suatu ucapan yang menyebabkan beberapa orang tersinggung [16]. Pengertian lain tentang Ujaran kebencian adalah aksi komunikasi yang di lakukan oleh orang ataupun kelompok

### 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.

dalam wujud provokasi, hasutan, maupun hinaan dengan tujuan untuk menjatuhkan orang ataupun kelompok lain dalam berbagai aspek seperti warna kulit, ras, etnis, gender, kecacatan, orientasi seksual, kewarganegaraan, agama dan lain-lain [6].

Menurut surat edaran Kapolri No.SE/06/X/2015 tentang Penanganan Ujaran Kebencian (*hatespeech*), Bahwa ujaran kebencian dapat berupa tindak pidana yang tercakup dalam Kitab Undang-Undang Hukum Pidana (KUHP) dan ketentuan hukum pidana lainnya selain KUHP yaitu berupa penghinaan, penodaan agama, penodaan agama, provokasi, hasutan, dan penyebaran berita bohong.

Ujaran kebencian (*hate speech*) dapat diungkapkan melalui berbagai media dan sarana, seperti kampanye, spanduk atau banner, media sosial, opini publik (demonstrasi), ceramah agama, media massa cetak dan elektronik, dll.

#### 2.1.3 *Abusive Language*

Bahasa kasar adalah ungkapan yang berisi kata-kata kasar, kotor dan cabul, baik lisan maupun tulisan [17]. Ujaran kebencian yang disertai bahasa kasar sering mempercepat terjadinya konflik sosial karena bahasa kasar dapat menyinggung perasaan dan memicu emosi [1]. Pada bahasa Indonesia bahasa kasar dapat diungkapkan salah satunya dengan menyebutkan spesies hewan tertentu, seperti monyet, anjing, dll, tetapi tidak semua kalimat yang mengandung kata hewan adalah bahasa kasar. Oleh karena itu, untuk menentukan apakah suatu kata dianggap sebagai kata kasar, perlu melihat keseluruhan konteks kalimat [18].

Mendeteksi suatu Bahasa kasar di media sosial tidak bisa dilakukan hanya dengan pencocokan kata. Karena sebagian besar pengguna jejaring sosial sering menggunakan ejaan dan tata bahasa informal. Mengklasifikasi bahasa kasar pada twitter lebih sulit untuk diselesaikan. karena pada data twitter, banyak pengguna yang memposting tweet dengan singkatan, karena terbatasnya jumlah kata yang diperbolehkan di twitter dalam sebuah postingan. [19]

#### 2.1.4 *Artificial Intelligence*

Menurut [20] *artificial intelligence* adalah membuat mesin dapat memiliki fungsi cognitive yang sama seperti manusia seperti belajar dan penyelesaian

### 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.

masalah. Menurut [20] definisi *Artificial intelligence* dibagi menjadi kategori yaitu *thinking humanly, thinking rationally, acting humanly, acting rationally*.

*Thinking humanly* : *artificial intelligence* merupakan upaya untuk membuat komputer berpikir seperti manusia (Haugeland ,1985).

*Thinking rationally* : *artificial intelligence* merupakan studi tentang membuat komputer dapat untuk melihat, bernalar, dan bertindak (Winston ,1992).

*Acting Humanly* : *artificial intelligence* merupakan seni menciptakan mesin yang dapat menyelesaikan tugas yang memerlukan kecerdasan yang di kontrol oleh manusia (Kurzweil, 1990).

*Acting Rationally* : *artificial intelligence* merupakan studi untuk mendesain agen cerdas (Poole et al., 1998).

#### 2.1.5 *Machine Learning*

Kata machine learning pertama kali dicetuskan oleh seorang ahli komputer dari Amerika Serikat, Arthur Samuel pada tahun 1959. menurut samuel *machine learning* adalah cabang studi ilmu komputer untuk membuat komputer dapat belajar untuk menyelesaikan suatu tugas tanpa diprogram. Menurut Mitchell *machine learning* adalah studi tentang algoritma komputer untuk meningkatkan algoritma secara otomatis melalui pengalaman dan pembelajaran. Machine learning merupakan bagian dari *artificial intelligence*. Model algoritma machine learning dibentuk melalui sampel data atau dikenal dengan data *training* yang digunakan untuk pengambilan keputusan atau prediksi [21].

Dalam pembelajaran machine learning, terdapat beberapa model seperti:

##### Supervised Learning

Supervised learning adalah model *machine learning* yang pembelajarannya menggunakan masukan data yang telah diberi label.

##### Unsupervised Learning

Unsupervide learning adalah model *machine learning* yang pembelajarannya menggunakan masukan data yang tidak memiliki label.

#### 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.

### 3 Semi-Supervised Learning

Merupakan gabungan dari supervised learning dan unsupervised learning. Pada model semi supervised, *dataset* untuk pelatihan sebagian memiliki label dan sebagiannya lagi tidak memiliki label.

### 4 Reinforcement Learning

Reinforcement Learning adalah model yang belajar menggunakan sistem reward dan penalties.

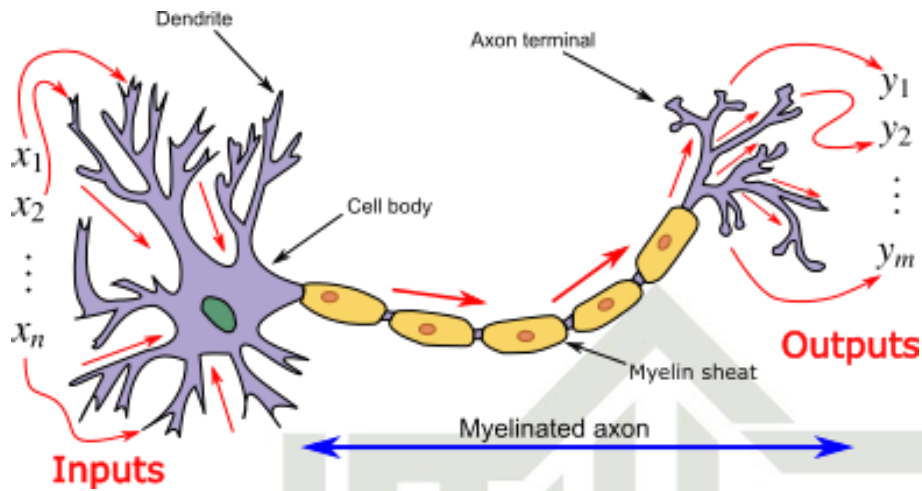
## 21.6 Artificial Neural Network (ANN)

*Artificial neural network* (jaringan syaraf tiruan) adalah sistem komputasi yang terinspirasi dari cara kerja jaringan syaraf yang ada pada otak manusia atau binatang [22]. Lapisan-lapisan penyusun ANN dibagi menjadi 3, yaitu lapisan input (*input layer*), lapisan tersembunyi (*hidden layer*), dan lapisan output (*output layer*). *Artificial neural network* (Jaringan syaraf tiruan) dibentuk berdasarkan kumpulan node-node atau disebut neuron, setiap node saling terhubung seperti sinapsis di otak manusia yang dapat mengirim sinyal ke neuron lainnya. Setiap neuron yang telah menerima sinyal akan memproses sinyal tersebut dan memberikannya ke neuron yang terhubung dengannya. Sinyal yang ada pada suatu koneksi adalah bilangan real dan *output* dari setiap neuron dihitung oleh fungsi non-linear. Penghubung antara node disebut *edge*. *Neuron* dan *edge* memiliki bobot yang dapat berubah tergantung proses pembelajaran yang dilakukan. Bobot dapat menambah atau mengurangi kekuatan sinyal pada sambungan. Neuron memiliki batas aktivasi sehingga sinyal dikirim hanya jika sinyal agregat melewati batas ambang (*threshold*).

Ide dasar jaringan saraf berasal dari otak manusia, dimana otak memiliki sekitar 86 miliar neuron. Neuron ini berfungsi untuk memproses semua informasi yang masuk. Sebuah neuron memiliki 1 akson dan setidaknya 1 dendrit. Setiap sel saraf terhubung ke saraf lain, dan masing-masing sel ini berinteraksi untuk menghasilkan kemampuan tertentu pada kerja otak manusia.

**Hak Cipta Dilindungi Undang-Undang**

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



**Gambar 1 Neuron pada Otak Manusia**

Dari gambar di atas, dapat dilihat bahwa ada beberapa bagian otak manusia, yaitu:

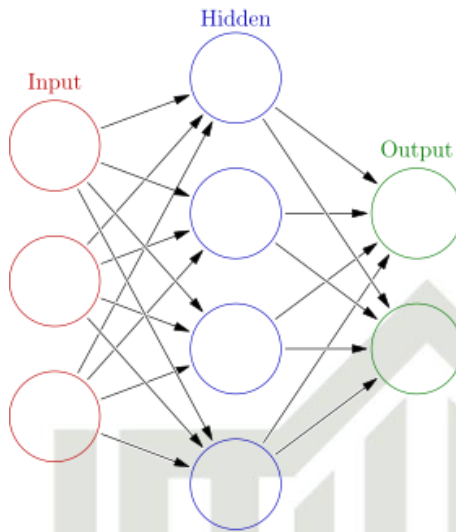
1. Dendrit (Dendrites) berfungsi untuk mengirimkan impuls yang diterima ke badan sel syaraf.
2. Akson (Axon) berfungsi untuk mengirimkan impuls dari badan sel ke jaringan lain
3. Sinapsis berfungsi sebagai unit fungsional di antara dua sel syaraf.

Proses yang terjadi pada otak manusia adalah neuron menerima impuls dari neuron lain melalui dendrit dan mengirimkan sinyal yang dihasilkan oleh badan sel melalui akson. Akson ini bercabang-cabang dan berkomunikasi dengan dendrit sel syaraf lain dengan mengirimkan impuls melalui sinapsis. Sinapsis adalah unit fungsional antara dua sel syaraf (misalnya, A dan B), satu adalah serat akson dari neuron A dan yang lainnya adalah dendrit dari neuron B. Kekuatan sinapsis bisa menurun atau meningkat tergantung besar penyiaran sinyal yang diterimanya. Impuls sinyal (informasi) akan diterima oleh neuron lain ketika berada dalam batas-batas tertentu yang biasa disebut dengan ambang batas (threshold).

Proses kerja *artificial neural network* memiliki konsep yang sama dengan proses kerja otak manusia.

**Hak Cipta Dilindungi Undang-Undang**

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



**Gambar 2 Struktur Artificial neural network**

Gambar diatas merupakan struktur *artificial neural network* yang menyerupai neuron pada otak manusia:

1. Input berfungsi sebagai dendrit
2. Output berfungsi seperti akson
3. Fungsi aktiva (*hidden layer*) berfungsi seperti sinapsis

Proses dalam jaringan syaraf tiruan dimulai dari input yang diterima oleh neuron beserta dengan nilai bobot dari setiap input. Setelah masuk ke dalam neuron, nilai input yang ada ditambahkan ke fungsi penjumlahan (*summing function*). Hasil penjumlahan diproses oleh fungsi aktivasi masing-masing neuron, proses ini akan membandingkan hasil penjumlahan dengan treshold tertentu. Jika nilainya lebih besar dari ambang batas, maka aktivasi neuron dibatalkan, sebaliknya jika di bawah ambang batas, maka neuron diaktifkan. Setelah neuron diaktifkan, ia akan mengirimkan nilai output melalui bobot-bobot outputnya ke semua neuron yang berhubungan dengannya. Proses ini akan terus berulang pada input-input selanjutnya [21].

### 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.

## 2.1.7 Deep Learning

*Deep Learning* adalah salah satu bagian dari Machine Learning berbasis *Artificial neural network* (ANN) atau bisa disebut sebagai perkembangan dari ANN. *Deep learning* merupakan metode *artificial neural network* yang terdiri dari beberapa *layer* untuk mempelajari representasi data dengan berbagai tingkat abstraksi. *layer* pada deep learning terdiri atas tiga bagian, yaitu *input layer*, *hidden layer*, dan *output layer*. Pada algoritma *deep learning*, *hidden layer* dibuat berlapis-lapis untuk menemukan komposisi algoritma yang tepat agar meminimalisir error pada output. Dalam deep learning, sebuah komputer belajar mengklasifikasi secara langsung dari gambar, teks, atau suara. Jenis pembelajaran dalam deep learning dapat berupa supervised, semi-supervised, dan unsupervised [23].

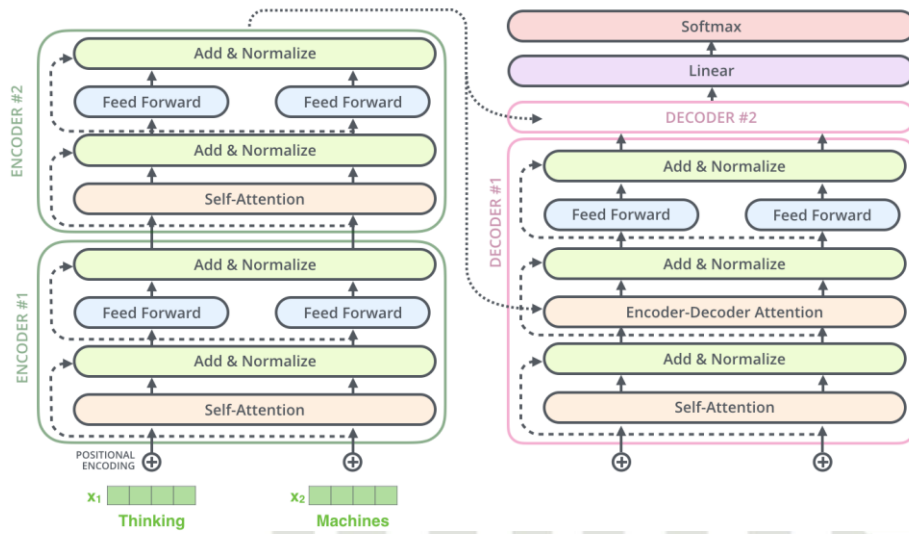
## 2.1.8 Transformer

*Transformer* adalah model *Deep learning* yang diperkenalkan pada tahun 2017 pada paper Attention is all you need [24], *transformer* digunakan dalam bidang *natural language processing*. *Transformer* dirancang untuk menangani data-data sekuensial, seperti *natural language processing* untuk tugas-tugas seperti penerjemahan dan peringkasan. *Transformer* adalah arsitektur encoder-decoder. Encoder terdiri dari sekumpulan lapisan encoding yang memproses *input* secara berulang-ulang satu demi satu dan decoder terdiri dari sekumpulan lapisan decoding yang melakukan hal yang sama pada *output* encoder. Fungsi dari setiap lapisan encoder adalah memproses masukannya untuk menghasilkan penyandian, berisi informasi tentang bagian mana dari masukan yang relevan satu sama lain kemudian meneruskan kumpulan penyandiannya ke lapisan pembuat encoder berikutnya sebagai masukan. Setiap lapisan decoder melakukan hal yang sebaliknya, mengambil semua pengkodean dan memprosesnya, menggunakan informasi kontekstual yang digabungkan untuk menghasilkan urutan keluaran [24].

Arsitektur *Transformer* dibangun berdasarkan lapisan-lapisan *Encoder* dan *decoder*.

**Hak Cipta Dilindungi Undang-Undang**

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



**Gambar 3 Arsitektur Encoder Decoder pada Transformer**

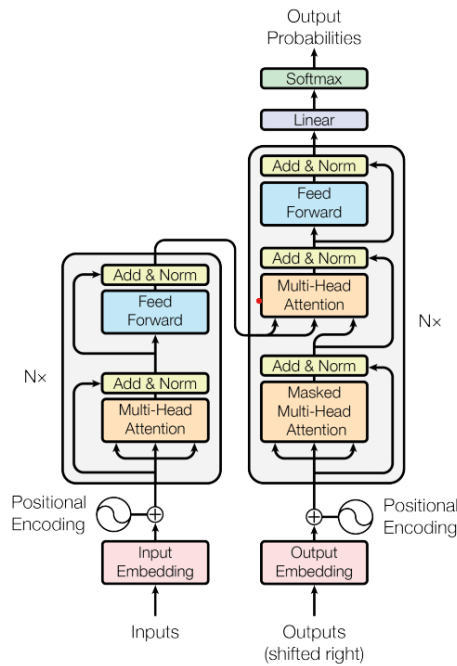
*Encoder* : encoder terdiri dari 6 layer. Setiap layer memiliki 2 sub-layers. layer pertama adalah multi-head self-attention mechanism, dan layer kedua adalah Position-wise fully connected feed-forward network. Diantara dua sub-layers tersebut terdapat layer normalisasi. Output dari setiap layer adalah  $\text{layerNorm}(x + \text{Sublayer}(x))$ .

*Decoder* : decoder terdiri dari 6 layer. Setiap layer memiliki 3 sub-layer. 2 sub-layer sama seperti yang ada pada encoder. Sub-layer ketiga terdiri dari multi-head attention over the output of the encoder stack. Sama seperti encoder setiap sub-layer terdapat layer normalisasi. Pada sub-layer ketiga terdapat masking untuk memastikan bahwa prediksi untuk posisi i hanya dapat bergantung pada output yang diketahui pada posisi kurang dari i.



**Hak Cipta Dilindungi Undang-Undang**

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

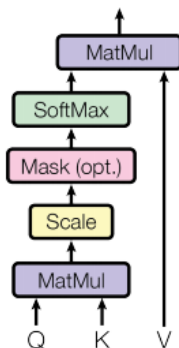


**Gambar 4 The Transformer - Model Architecture**

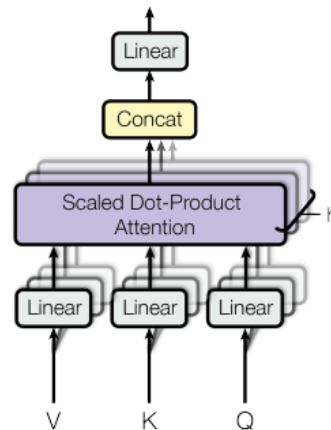
1. *Attention*

Fungsi *attention* dapat dideskripsikan sebagai pemetaan query dan sekumpulan pasangan key-value ke output, di mana *query*, *key*, *value*, dan output semuanya adalah vektor. Outputnya dihitung dari total bobot dari value, di mana bobot yang ditetapkan ke setiap nilai dihitung oleh fungsi kompatibilitas dari *query* dengan *key* yang sesuai.

Scaled Dot-Product Attention



Multi-Head Attention



**Gambar 5 Scaled Dot-Product Attention dan Multi-Head Attention**

**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.

a. *Scaled Dot-Product Attention*

*scaled dot-product attention* terdiri dari *query* dan *key* dimensi  $d_k$ , dan *value* dimensi  $d_v$ . untuk menghitung *scaled dot-product*, *query* dikalikan dengan semua *key*, dan dibagi dengan  $\sqrt{d_k}$ , dan terapkan fungsi softmax untuk mendapatkan bobot pada *value*.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{D_k}}\right)V$$

b. *Multi-Head Attention*

*Multi-head attention* merupakan kumpulan fungsi *scaled dot-product attention*. *Multi-head attention* memungkinkan model untuk bersama-sama memperhatikan informasi dari representasi yang berbeda diposisi berbeda.

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O$$

$$\text{head}_i = \text{Attention}(QW^Q, KW^K, VW^V)$$

2. *Position-wise Feed-Forward Networks*

setiap *layer* encoder dan decoder terdapat *fully connected feed-forward network* yang terdiri dari 2 *linear transformations* dengan aktivasi RELU.

$$\text{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2$$

**21.9 Bidirectional Encoder Representations from Transformers (BERT)**

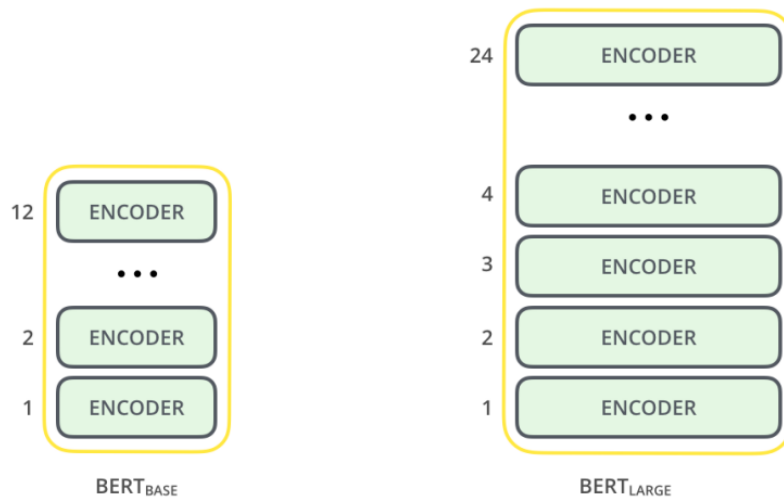
*Bidirectional encoder representations from transformers* (BERT) adalah teknik *machine learning* berbasis *Transformer* untuk *pre-training* pemrosesan bahasa alami (NLP) yang dikembangkan oleh Google. BERT dibuat dan diterbitkan pada 2018 oleh Jacob Devlin dan rekan-rekannya dari Google Pada 2019, Google telah memanfaatkan BERT untuk lebih memahami penelusuran pengguna. Sebelumnya dalam representasi word *embedding* seperti GloVe, word2vec dan FastText, kata seperti “tahu” memiliki makna yang sama pada kalimat “saya tidak tahu” dengan “saya suka makan tahu”. untuk menyelesaikan permasalahan tersebut muncul ELMo (bi-directional LSTM) yang menggunakan kata sebelumnya dan sesudahnya untuk menghasilkan *embedding*. Kemudian muncul BERT yang

**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.

memiliki performa lebih baik dari model sebelumnya. BERT dirancang untuk melatih representasi dua arah dari teks yang tidak berlabel dengan mengkondisikan konteks kiri dan kanan setiap kata di semua lapisan. Hasilnya, model BERT yang telah dilatih sebelumnya dapat dilakukan *fine-tuning* dengan menambah satu output *layer* tambahan untuk membuat state-of-the-art model untuk berbagai tugas natural language processing seperti *teks classification*, *question answering* dan *translation*. [11].

BERT model architecture dibentuk dari *multilayer* bi-directional Transformer encoder. terdapat 2 bentuk model :



**Gambar 6 Arsitektur BERT**

1. BERT BASE : model ini dibangun dari 12 Transformer block, 12 Attention layer dan 768 hidden layer.
2. BERT LARGE : model ini memiliki layer dan attention layer yang lebih banyak dari BERT BASE untuk mendapatkan hasil yang lebih baik yaitu 24 transformer block, 16 attention head dan 1024 hidden layer.

BERT dapat digunakan untuk *pre-training* dan *fine-tuning*, pada *pre-training*, model dilatih menggunakan data yang tidak berlabel untuk menyelesaikan berbagai tugas *pre-training*. Pada *fine-tuning*, model BERT diinisialisasi dengan

### 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.

parameter yang telah dilatih sebelumnya dan semua parameter tersebut di *fine-tuning* dengan data yang memiliki label untuk menyelesaikan berbagai tugas nlp seperti teks klasifikasi, *question-answering*, *Named entity recognition*.

#### 1. *Pre-training*

*pre-training* BERT dapat digunakan untuk menyelesaikan tugas unsupervised seperti *Masked model language* dan *next sentence prediction*.

##### a. *Masked language model*

*Masked language model* adalah model untuk memprediksi kata yang seharusnya terdapat pada sebuah kalimat. Contoh: input: saya sudah menonton [MASK] itu. output saya sudah menonton film itu. Sebelum memasukkan kata ke BERT, 15% kata di setiap urutan diganti dengan *token* [MASK]. Model kemudian mencoba untuk memprediksi nilai asli dari kata mask tersebut berdasarkan konteks yang diberikan oleh kata lain dalam urutan.

##### b. *Next sentence prediction*

*Next sentence prediction* adalah model yang dapat memprediksi kata selanjutnya dari sebuah kalimat. Dalam proses pelatihan BERT, model menerima pasangan kalimat sebagai masukan dan belajar memprediksi kalimat kedua pada pasangan tersebut jika kalimat berikutnya dalam dokumen aslinya. Selama pelatihan, 50% masukan berupa pair dimana kalimat kedua merupakan kalimat berikutnya pada dokumen asli, sedangkan pada 50% lainnya dipilih kalimat acak dari korpus sebagai kalimat kedua. Asumsinya adalah bahwa kalimat acak akan terputus dari kalimat pertama.

#### 2. *Fine-tuning*

Model *pre-training* BERT yang telah dilatih dapat digunakan untuk menyelesaikan tugas *nlp* lainnya dengan menambahkan *layer* yang sesuai dengan tugas yang ingin di selesaikan pada model yang ada. BERT dapat

digunakan untuk klasifikasi teks, sentiment analysis, question-answering, named entity recognition dan tugas *natural language* lainnya.

### 2.1.10 Transfer Learning

*Transfer learning* adalah penelitian *machine learning* yang berfokus pada penyimpanan pengetahuan yang diperoleh dari data latih yang digunakan untuk menyelesaikan masalah tertentu dan dapat digunakan untuk menyelesaikan masalah lainnya yang berkaitan. Contohnya pengetahuan yang didapatkan saat belajar untuk mengenali mobil dapat digunakan untuk mengenali truk. *transfer learning* biasanya digunakan untuk tugas-tugas yang memiliki sedikit data *training* [11].

*Transfer learning workflow* :

1. Ambil *layer* dari model yang dilatih sebelumnya.
2. bekukan *layer* tersebut untuk menghindari rusaknya informasi yang ada pada model sebelumnya.
3. Tambahkan beberapa *layer* baru yang bisa dilatih di atas *layer* yang telah dibekukan. *layer* tersebut akan belajar mengubah fitur lama menjadi prediksi pada kumpulan data baru.
4. Latih *layer* baru pada *dataset*.

### 2.1.11 Confusion Matrix

*Confusion Matrix* adalah suatu metode pengujian untuk menghitung tingkat ketepatan suatu model untuk menentukan benar atau salah. Pada *confusion matrix* dapat dihitung akurasi, *precision* dan *recall*. Nilai akurasi adalah tingkat prediksi data benar dengan keseluruhan data. *Precision* adalah tingkat ketepatan antara informasi yang diminta dengan jawaban yang diberikan oleh sistem. *Recall* adalah tingkat keberhasilan sistem dalam menemukan kembali informasi [25]. Berikut merupakan tabel dari *confusion matrix*.

#### 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.



**Tabel 1 Confusion Matrix**

	Actual Values	
Predicted Values	True Negative (TN)	False Positif (FP)
	False Negative (FN)	True Positif (TP)

$$\text{Accuracy} = \frac{TP+TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

Keterangan:

TP : True Positive merupakan data yang sebenarnya bernilai positif dan hasil klasifikasinya juga bernilai positif

TN : True Negative merupakan data yang sebenarnya bernilai negatif dan hasil klasifikasinya juga bernilai negatif

FP : False Positive merupakan data yang sebenarnya bernilai positif dan hasil klasifikasinya bernilai negatif

FN : False Negative merupakan data yang sebenarnya bernilai negatif dan hasil klasifikasinya bernilai positif

## 2.2 Penelitian Terkait

Tabel berikut ini merupakan ringkasan dari penelitian sebelumnya yang berkaitan dengan penelitian tugas akhir ini.

**Tabel 2 Penelitian Terkait**

No	Penulis	Judul	Tahun	Kesimpulan
1	Marzieh Mozafari, Reza	A BERT-Based Transfer Learning Approach for	2019	Pada penelitian ini menggunakan pre-trained BERT yang dilatih dari

- 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. 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.

	Farahbakhsh, dan No'el Crespi [12]	Hate Speech Detection in Online Social Media		Wikipedia berbahasa inggris dan bookcorpus model BERT mendapatkan akurasi sebesar 92% dengan <i>fine-tuning</i> terbaik BERT+CNN.
	Cindy Alifia Putri, Adiwijaya, Said Al Faraby [14]	Analisis Sentimen Review Film Berbahasa Inggris dengan Pendekatan Bidirectional Encoder Representations from Transformers	2020	Penelitian ini menggunakan 2000 data. 1000 data berlabel positif dan 1000 berlabel negatif. Pada penelitian peneliti pertama menggunakan algoritma naive bayes dan mendapatkan akurasi 48% kemudian menggunakan metode BERT dan mendapatkan akurasi 78%
3	David Junggu Manggala Pasaribu, Kusri, Sudarmawan [26]	Peningkatan Akurasi Klasifikasi Sentimen Ulasan Makanan Amazon Dengan Bidirectional LSTM dan Bert <i>embedding</i>	2020	Pada penelitian ini penulis membandingkan penelitian sebelumnya tentang sentimen ulasan makanan amazon dengan menggunakan berbagai metode machine learning dan deep learning. Pada penelitian sebelumnya algoritma LSTM mendapatkan akurasi 90% dan pada penelitian ini menggunakan metode BERT mendapatkan akurasi lebih tinggi yaitu 92%.
	Muhammad Okky Ibrohim, Indra Budi [1]	Multi-label Hate Speech and <i>abusive language</i> Detection in Indonesian Twitter	2019	Pada penelitian ini membandingkan beberapa metode machine learning seperti suport vector machine, naive bayes dan random forest decision tree dan menggunakan beberapa teknik feature extraction. Dari penelitian tersebut didapatkan bahwa algoritma random forest decision tree



				mendapatkan akurasi tertinggi yaitu 76,6%
	Ahmad Fathan Hidayatullah, Afa Aulia Fadila, Kiki Purnama, Royan Abida [18]	Identifikasi Konten Kasar pada Tweet Bahasa Indonesia	2019	Penelitian ini menggunakan 5462 tweet, 2731 berlabel kata kasar dan 2731 tidak kasar. Pada penelitian ini menggunakan metode naive bayes dan support vector machine.
	Junanda Patihullah, Edi Winarko [4]	Hate Speech Detection for Indonesia Tweets Using Word <i>embedding</i> And Gated Recurrent Unit	2019	Penelitian ini menggunakan algoritma deep learning Gated Recurrent Unit dan menggunakan feature extraction word2vec dan tf-idf dan membandingkannya dengan metode svm, naive bayes, random forest, regresi logistik. Hasil yang didapatkan menunjukkan Gated Recurrent Unit dengan word2vec mendapatkan hasil terbaik yaitu 92,96%
7	Ibrahim Abu-Farha dan Walid Magdy [27]	Multitask Learning for Arabic Offensive <i>language</i> and Hate-Speech Detection	2020	Pada penelitian ini menggunakan beberapa model deep learning multitask learning dan transfer learning. Penelitian ini menggunakan model Bidirectional-LSTM dan CNN-BiLSTM. Hasilnya metode CNN-BiLSTM mendapatkan akurasi terbaik yaitu 90,1 %
	Muhammad Okky Ibrohim, Muhammad Akbar Setiadi, Indra Budi [7]	Identification of Hate Speech and <i>abusive language</i> on Indonesian Twitter Using theWord2vec, Part of Speech and Emoji Features	2019	Pada penelitian ini menggunakan word2vec sebagai feature extraction dan dikombinasikan dengan part of speech dan emoji feature. kombinasi word unigram,

**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

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.

				character quadgrams, positive sentiment lexicon, and <i>abusive</i> lexicon dengan algoritma random forest decision tree mendapatkan 77.36% akurasi.
	Son T. Luu, Hung P. Nguyen, Kiet Van Nguyen, dan Ngan Luu-Thuy Nguyen [28]	Comparison Between Traditional Machine Learning Models And Neural Network Models For Vietnamese Hate Speech Detection	2020	Pada penelitian ini menggunakan 20345 <i>dataset</i> . Penelitian ini membandingkan algoritma machine learning (SVM, Logistic Regresion) dan deep learning (GRU, CNN). hasil penelitian ini menunjukkan algoritma deep learning memiliki akurasi yang lebih baik dari algoritma machine learning tradisional yaitu CNN mendapatkan akurasi 83.04% sedangkan SVM 65.10%.
10	Pinkesh Badjatiya, Shashank Gupta, Manish Gupta, Vasudeva Varma1 [29]	Deep Learning for Hate Speech Detection in Tweets	2017	Penelitian ini membandingkan beberapa algoritma deep learning seperti CNN dan LSTM menggunakan beberapa fitur extraction seperti n-gram, TF-IDF, bag of word, Glove, Fasteks. Dari penelitian ini didapatkan bahwa algoritma LSTM dikombinasikan dengan random <i>embedding</i> mendapatkan akurasi 93%
11	Dayang Putri Nur Lyrawati [30]	DETEKSI UJARAN KEBENCIAN PADA TWITTER MENJELANG PILPRES 2019 DENGAN MACHINE	2019	Pada penelitian ini, deteksi ujaran kebencian pada data twitter menjelang pemilihan presiden 2019 dengan machine learning menggunakan algoritme



**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.

		LEARNING		SVM-RBF Kernel. Parameter yang digunakan pada algoritme SVM-RBF Kernel adalah C dengan nilai 1, 5, dan 10. Berdasarkan penelitian yang telah dilakukan, dihasilkan akurasi tertinggi sebesar 61,667% dicapai saat rasio 9:1 dengan ketiga nilai epsilon yang berbeda dengan C=1
2	Ibrahim Abu-Farha dan Walid Magdy [27]	Multitask Learning for Arabic Offensive language and Hate-Speech Detection	2020	Pada penelitian ini melakukan beerapa <i>experiment</i> dengan menggunakan deep learning, transfer learning dan multi tasklearning. Hasil penelitian ini mendapatkan model terbaik yaitu multi task learning dengan algoritma CNN-BiLSTM dengan akurasi 90%

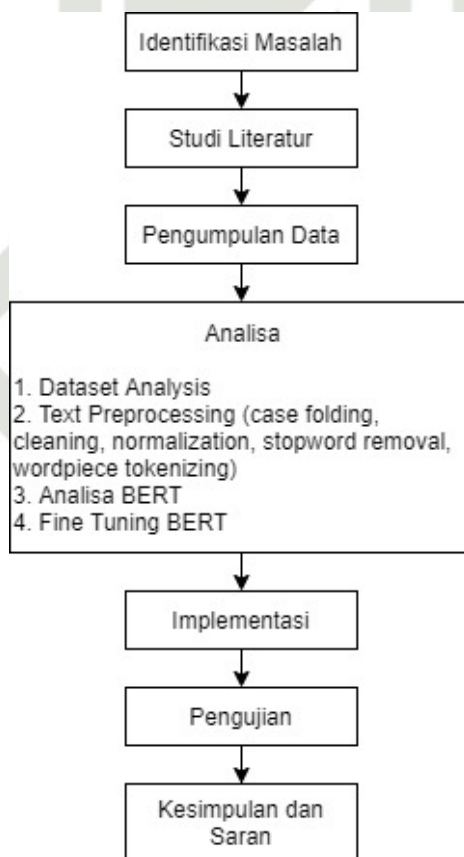
**Hak Cipta Dilindungi Undang-Undang**

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.

## BAB 3

### METODOLOGI PENELITIAN

Metodologi penelitian merupakan pedoman atau tahapan-tahapan dalam melakukan penelitian. Metodologi penelitian ini dilakukan bertujuan untuk mendapatkan hasil yang sesuai dengan yang diinginkan. Berikut tahapan-tahapan dalam penelitian:



**Gambar 7 Tahapan Metodologi Penelitian**

#### 3.1 Identifikasi Masalah

Pada tahap ini dilakukan analisa tentang permasalahan yang akan diselesaikan dengan menentukan latar belakang, rumusan masalah, tujuan dari penelitian dan batasan-batasan yang pada penelitian ini. Adapun hal-hal yang dilakukan pada tahap identifikasi masalah yaitu sebagai berikut:

**Hak Cipta Dilindungi Undang-Undang**

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

1. Latar Belakang

Latar belakang merupakan landasan atau dasar penelitian ini perlu dilakukan, dalam penelitian ini fenomena *hate speech* dan *abusive language* pada media sosial yang dapat menimbulkan konflik atau perpecahan dimasyarakat.

2. Tujuan Penelitian

Tujuan penelitian merupakan hal-hal yang ingin dicapai dalam penelitian. Pada penelitian ini tujuan utamanya yaitu untuk mengklasifikasikan tweet apakah tergolong ke dalam *hate speech* atau *abusive* dan *level hate speech* dari tweet tersebut apakah kuat, sedang, atau lemah.

3. Rumusan Masalah

Rumusan masalah merupakan masalah-masalah akan diselesaikan dalam penelitian.

4. Batasan Masalah

Batasan masalah merupakan batasan-batasan yang dilakukan pada penelitian agar penelitian lebih fokus dan dapat tercapai.

**3.2 Studi Literatur**

Pada tahap ini dilakukan pencarian informasi, teori dan konsep-konsep dasar mengenai materi yang berhubungan dengan penelitian. Pencarian informasi ini didapat dengan cara mempelajari buku-buku, artikel, jurnal, skripsi, video youtube dan referensi lainnya yang berhubungan dengan penelitian.

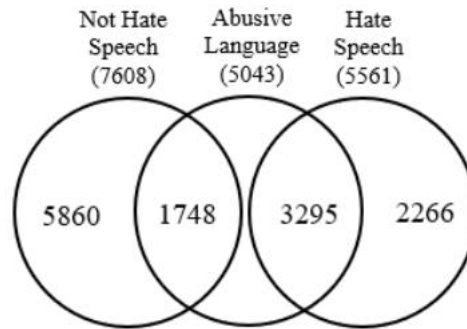
**3.3 Pengumpulan Data**

Penelitian ini menggunakan *dataset tweet hate speech* dan *abusive language* berbahasa indonesia dari beberapa penelitian sebelumnya yang terdiri dari [31], [32], dan [33]. *Dataset* pada penelitian ini berjumlah 13.169 tweet. *Dataset* pada penelitian ini merupakan dataset dengan data multi-label, jadi pada setiap data terdapat beberapa label. Pada dataset ini data bisa saja berlabel hatespeech tetapi bukan abusive bisa juga abusive dan bukan hatespeech dan bisa juga keduanya atau tidak keduanya. *Dataset* ini terdiri dari 7.608 tweets tidak *hatespeech*, 5.561 tweet

**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.

*hatespeech* dan terdapat 5.043 tweet *abusive language*. Berikut gambar pembagian *dataset* pada penelitian ini.



**Gambar 8 Pembagian Dataset Hatespeech dan Abusive Language**

Dari dataset tersebut terdapat 12 label yaitu *hatespeech* 5.561 tweet, *abusive* 5.043 tweet, target *hatespeech* (*individual* 3.575 tweet, *group* 1.986 tweet), *categories hatespeech* (*religion/creed* 793 tweet, *race/ethnicity* 566 tweet, *physical/disability* 323 tweet, *gender/sexual* 306 tweet, other *invective/slander* 3.740 tweet) dan level (*weak* 3.383 tweet, *moderate* 1.704 tweet, *strong* 473 tweet). Berikut penjelasan label dari dataset.

1. *Hatespeech*  
Ujaran kebencian.
2. *Abusive Language*  
Bahasa kasar.
3. *Individual hatespeech*  
*Individual hatespeech* merupakan ujaran kebencian yang ditujukan pada individu tertentu.
4. *Group hatespeech*  
*Group hatespeech* merupakan ujaran kebencian yang ditujukan kepada suatu kelompok tertentu.
5. *Religion/creed hatespeech*  
Ujaran kebencian berdasarkan pada suatu agama.

**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.

6. *Race/ethnicity hatespeech*

Ujaran kebencian berdasarkan ras manusia (kelompok manusia berdasarkan ciri fisik seperti bentuk wajah, tinggi badan, warna kulit dan lain-lain) atau suku (kelompok manusia berdasarkan kewarganegaraan atau tradisi budaya di suatu wilayah)

7. *Physical/disability hatespeech*

Ujaran kebencian berdasarkan kekurangan/perbedaan fisik (Misalnya bentuk wajah, mata, dan tubuh bagian lainnya) atau cacat (misalnya autisme, idiot, buta, tuli, dll).

8. *Gender/sexual hatespeech*

Ujaran kebencian berdasarkan jenis kelamin (laki-laki dan perempuan), mengutuk seseorang (atau kelompok) menggunakan kata-kata yang merendahkan gender (misalnya: gigolo, jalang, dll.), atau orientasi seksual yang menyimpang (misalnya: homoseksual, lesbian, dll.);

9. *Invective/slander hatespeech*

Ujaran kebencian berupa makian/ejekan dengan menggunakan kata-kata kasar kata/frasa atau fitnah/hasutan lainnya yang tidak berhubungan dengan keempat kelompok yang telah dijelaskan sebelumnya.

10. *Hatespeech weak*

Hatespeech weak yaitu ujaran kebencian dalam bentuk makian/fitnah yang ditujukan pada individu tanpa menyertakan hasutan/provokasi untuk memunculkan konflik terbuka. Di Indonesia, ujaran kebencian dalam bentuk ini dikategorikan sebagai ujaran kebencian yang lemah karena merupakan masalah pribadi. Artinya, jika korban ujaran kebencian tidak lapor ke pihak berwajib (merasa biasa dan pemaaf) maka ujaran kebencian itu tidak terlalu diprioritaskan untuk diselesaikan oleh otoritas.

11. *Hatespeech moderate*

Hatespeech moderate yaitu ujaran kebencian dalam bentuk sumpah serapah/penistaan agama/stereotyping/labeling yang ditujukan kepada kelompok-kelompok tanpa melakukan provokasi untuk memunculkan konflik terbuka. Meskipun dapat mengundang konflik antar kelompok,

ujaran kebencian semacam ini diperkirakan hanya terbatas pada konflik di media sosial.

#### 12. *Hatespeech strong*

Hatespeech strong yaitu ujaran kebencian dalam bentuk makian/fitnah/penistaan agama/stereotyping/labeling yang ditujukan kepada individu atau kelompok termasuk hasutan/provokasi untuk memunculkan konflik terbuka. ujaran kebencian seperti ini perlu diprioritaskan untuk diselesaikan segera karena dapat mengundang konflik yang meluas dan dapat menimbulkan konflik/fisik kehancuran di dunia nyata.

Proses anotasi pada *dataset* ini dilakukan dengan cara *crowdsourcing* berbayar [1]. Pada *dataset* ini proses anotasi dilakukan oleh 30 anotator dari berbagai daerah di Indonesia. Terdapat 14 laki-laki anotator dan 16 perempuan anotator. Anotator berasal dari berbagai macam pekerjaan, suku, dan agama. pekerjaan anotator terdiri dari mahasiswa sarjana (12 anotator), mahasiswa master (3 anotator), PNS (1 anotator), pegawai honorer (1 anotator), guru / tutor / asisten pengajar (5 anotator), dan karyawan swasta (8 anotator). Etnis asal anotator terdiri dari Jawa (11 anotator), Bali (4 anotator), Tionghoa (4 anotator), Betawi (3 anotator), Batak (2 anotator), dan lain-lain (6 anotator, berasal dari Melayu, Minang, Sunda, Cirebon, Ambon, Toraja) dan Agama anotator terdiri dari Islam (5 anotator), Kristen (5 anotator), Katolik (5 anotator), Hindu (3 anotator), dan Buddha (2 anotator).

### 3.4 Analisa

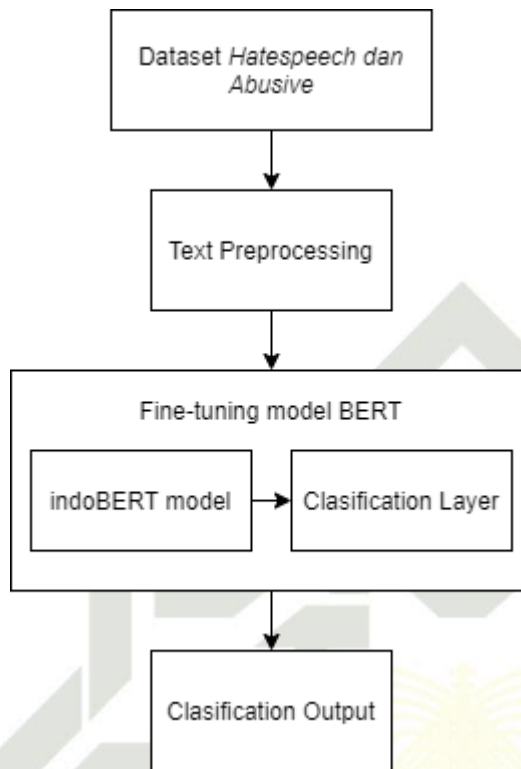
Setelah dilakukannya identifikasi masalah, studi literatur dan pengumpulan data selanjutnya dilakukan tahapan analisa. Analisa merupakan tahapan untuk mempelajari dan melakukan evaluasi terhadap suatu permasalahan serta bertujuan untuk mengetahui gambaran jelas mengenai penelitian yang dilakukan. Tahapan analisa yang dilakukan sebagai berikut :

#### Hak Cipta Dilindungi Undang-Undang

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

**Hak Cipta Dilindungi Undang-Undang**

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



**Gambar 9 Tahapan Analisa Model BERT**

### 3.4.1 Analisa Dataset

Pada tahap ini dilakukan persiapan terhadap data yang akan digunakan dan menentukan apakah data sudah baik dan tidak terdapat data yang bernilai kosong. Pada tahap ini juga dilakukan analisa terhadap data dan melihat mana data yang akan mempengaruhi akurasi pada penelitian. Pada penelitian ini *dataset* yang didapat akan dilakukan pembagian data latih dan data uji. Pembagian data latih dan data uji yang akan dilakukan yaitu 80:20 dan 90:10. Pembagian data latih dan data uji dilakukan agar pada saat pengujian terdapat data yang terpisah yang akan digunakan untuk menguji model.

### 3.4.2 Teks Preprocessing

*Teks preprocessing* merupakan langkah yang bertujuan untuk membersihkan data dari unsur-unsur yang tidak dibutuhkan dan dapat digunakan pada proses klasifikasi.



#### 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.

Tahapan *preprocessing* yang dilakukan pada penelitian ini:

a. *Case folding*

*Case folding* merupakan proses untuk mengubah seluruh huruf di dalam dokumen menjadi huruf kecil.

b. *Cleaning*

*Cleaning* merupakan proses untuk membersihkan dokumen dari komponen - komponen yang tidak memiliki hubungan dengan informasi yang ada pada dokumen. pada *dataset* ini akan dihapus komponen yang kurang penting seperti kata user, rt, *url*, *emoji*, angka dan tanda-tanda baca.

c. *Normalization*

*Normanization* merupakan proses untuk merubah kata yang salah ketik atau kata yang tidak baku pada dataset menjadi kata baku.

d. *Stopword removal*

*Stopword removal* merupakan tahapan untuk menghapus *stopword* atau kata-kata sering muncul yang dianggap kurang penting

e. *Wordpiece tokenizing*

*Wordpiece tokenizing* merupakan tahapan yang dilakukan untuk memisahkan kalimat menjadi kata dan imbuhan pada kata. Pemisahan tersebut tanpa melalui proses *stemming* dan *lemmatizing*. Pada umumnya jika menggunakan penghasil vektor kata biasa (*General Word embedding*) seperti *Word2Vec*, *fastteks* dan *Gloove* tahap pra proses tersebut penting dilakukan karena untuk mengurangi data yang tidak memiliki makna untuk diproses oleh komputer namun pada peneliitan ini menggunakan penghasil vektor kata secara kontekstual (*Conteksual Word embedding*) menggunakan *BERT*, dimana kata-kata tersebut penting untuk mencari vektor setiap kata yang sesuai dengan konteks kalimat.

### 3.4.3 Analisa BERT

Pada tahap ini dilakukan analisa pada pretrain *BERT* yang akan digunakan pada penelitian ini. Model *pre-trained* *BERT* yang digunakan merupakan model *indoBERT* yang di dapat dari penelitian [13]. Model *indoBERT* ini di latih dari

### 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.

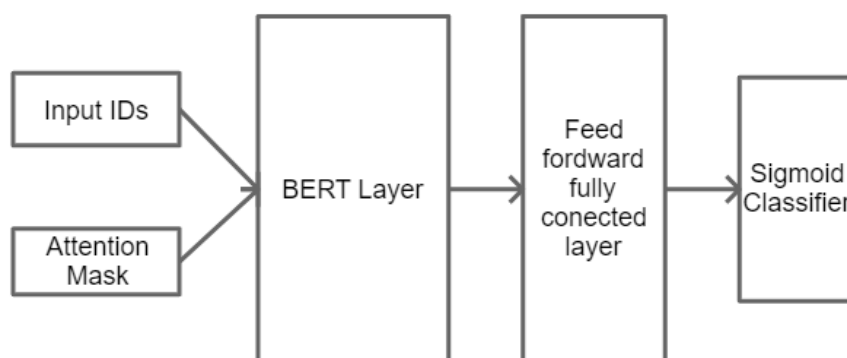
*dataset* indo4B. *dataset* indo4B merupakan *dataset* bahasa indonesia yang di dapat dari sumber publik seperti sosial media, blog, berita dan website.

#### 3.4.4 *Fine tuning BERT*

Pada tahap ini model pretrain BERT yang telah di dapatkan akan dilakukan *fine-tuning* yaitu dengan menambahkan *layer* yang digunakan untuk proses klasifikasi. Pada penelitian ini akan ditambahkan *fully connected neural network layer* untuk mengklasifikasi teks *hatespeech* dan *abusive*.

Berikut tahapan *fine-tuning* BERT:

1. Input IDs dan *attention mask* yang didapat setelah proses *tokenisasi* digunakan sebagai inputan pada BERT.
2. Input IDs dan *attention mask* diproses pada BERT *layer*. Output dari proses ini adalah vektor.
3. Output dari proses BERT *layer* selanjutnya diproses pada neural network menggunakan *feed forward fully conected layer*.
4. Setelah proses pada *neural network* maka akan dihasilkan bobot yang selanjutnya dilakukan proses klasifikasi dengan metode sigmoid.
5. Hasil dari proses klasifikasi merupakan label data.



Gambar 10 Tahapan *fine-tuning* BERT

**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.

### 3.5 Implementasi

Implementasi adalah proses penerapan model berdasarkan hasil dari analisa dan perancangan. Proses ini membutuhkan perangkat pendukung berupa perangkat keras dan perangkat lunak.

1. Perangkat keras

Perangkat keras yang akan digunakan dalam penelitian ini yaitu Laptop dengan Processor Intel Core i5, RAM 8 GB dengan sistem operasi windows.

2. Perangkat lunak

Perangkat lunak yang akan digunakan dalam penelitian ini yaitu menggunakan python, jupyter notebook.

### 3.6 Pengujian

Pada tahap ini dilakukan pengujian kinerja dari suatu model yang telah dibangun dengan menghitung tingkat akurasi dari model dengan menggunakan *confusion matrix*. Pada tahap ini model diterapkan berdasarkan perancangan skenario eksperimen. Eksperimen yang akan dilakukan:

1. Pembagian data latih dan uji 80:20 dan 90:10.
2. Melakukan eksperimen tahapan *preprocessing* dengan melakukan kombinasi penggunaan preprocessing.
3. Berbagai percobaan nilai hyperparameter seperti *learning rate* dan *epoch*.

### 3.7 Kesimpulan dan Saran

Pada tahap ini merupakan tahap terakhir dari penelitian ini. Pada tahap ini akan dijelaskan mengenai hasil dan kesimpulan dari penelitian yang telah dilakukan yaitu implementasi BERT untuk mengklasifikasi *hatespeech* dan *abusive* pada twitter bahasa indonesia dan pada tahap ini juga akan dijelaskan saran yang perlu dilakukan oleh penelitian berikutnya.



## BAB 5 PENUTUP

### 5.1 Kesimpulan

Berdasarkan hasil pengujian dan beberapa *experiment* yang dilakukan, maka dapat ditarik beberapa kesimpulan yaitu :

1. Model klasifikasi BERT ini dapat diterapkan untuk proses klasifikasi *multi-label hatespeech* dan *abusive language* pada twitter bahasa indonesia.
2. Model klasifikasi BERT mendapatkan hasil akurasi yang lebih baik dari penelitian sebelumnya [1] dalam melakukan klasifikasi *hatespeech* dan *abusive language* beserta target, kategori dan level *hatespeech* yaitu **72.28%** dan untuk mengklasifikasi *hatespeech* dan *abusive language* tanpa level, target dan kategori adalah **85.80%**.
3. Kombinasi *hyperparameter* terbaik pada model klasifikasi BERT yang telah dibangun pada penelitian ini adalah *epoch=3, learning rate=0.00001*.
4. Pada model klasifikasi BERT yang telah dibangun penggunaan teks *preprocessing (case folding, cleaning, normalization dan stopword removal)* tidak meningkatkan hasil akurasi.
5. Pada hasil pengujian menggunakan *confusion matrix* diketahui bahwa model yang telah dibangun memiliki performa yang kurang baik dalam mengklasifikasi label yang ditraining menggunakan data yang sedikit.

### 5.2 Saran

Saran yang dapat diberikan untuk digunakan pada penelitian berikutnya adalah sebagai berikut :

1. Peneliti berikutnya dapat menggunakan *pretaining language* model lainnya seperti *Universal language Model fine-tuning (ULMFiT), embedding from language Models (ELMO), OpenAI Generative Pre-trained Transformer (GPT)*.

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.



2. Penelitian berikutnya dapat menggunakan klasifikasi *deep learning* lainnya seperti CNN, LSTM dan RNN.
3. Penelitian berikutnya dapat menambahkan dataset yang lebih banyak pada label yang tidak seimbang.



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, 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. 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 PUSTAKA

- M. O. Ibrohim and I. Budi, "Multi-label Hate Speech and Abusive Language Detection in Indonesian Twitter," pp. 46-57, 2019.
- H. Juwiantho, E. I. Setiawan, J. Santoso and M. H. Purnomo, "Sentiment Analysis Twitter Bahasa Indonesia Berbasis Word2Vec Menggunakan Deep Convolutional Neural Network," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 7, no. 1, pp. 181-188, 2020.
- We Are Social & Hootsuite, "Indonesia Digital report 2020," *Global Digital Insights*, pp. 1-92, 2020.
- [4] J. Patihullah and E. Winarko, "Hate Speech Detection for Indonesia Tweets Using Word Embedding And Gated Recurrent Unit," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 13, no. 1, p. 43, 2019.
- [5] A. Al-Hassan and H. Al-Dossari, "Detection of Hate Speech in Social Networks: a Survey on Multilingual Corpus," pp. 83-100, 2019.
- [6] A. Sutantohadi, "Bahaya Berita Hoax Dan Ujaran Kebencian Pada Media Sosial Terhadap Toleransi Bermasyarakat," *DIKEMAS (Jurnal Pengabdian Kepada Masyarakat)*, vol. 1, no. 1, pp. 1-5, 2018.
- [7] M. O. Ibrohim, M. A. Setiadi and I. Budi, "Identification of hate speech and abusive language on Indonesian twitter using theword2vec, part of speech and emoji features," *ACM International Conference Proceeding Series*, 2019.
- [8] M. Teja, "Media Sosial: Ujaran Kebencian dan Persekusi," *Majalah Info Singkat*, vol. IX, no. 11, pp. 9-12, 2017.



**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.

[9] M. S. Saputri, R. Mahendra and M. Adriani, "Emotion Classification on Indonesian Twitter Dataset," *Proceedings of the 2018 International Conference on Asian Language Processing, IALP 2018*, pp. 90-95, 1 2019.

[10] M. N. Fakhruzzaman, S. Z. Jannah, R. A. Ningrum and I. Fahmiyah, "Clickbait Headline Detection in Indonesian News Sites using Multilingual Bidirectional Encoder Representations from Transformers (M-BERT)," 2021.

[11] J. Devlin, M. W. Chang, K. Lee and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, 2019.

[12] M. Mozafari, R. Farahbakhsh and N. Crespi, "A BERT-Based Transfer Learning Approach for Hate Speech Detection in Online Social Media," *Studies in Computational Intelligence*, vol. 881 SCI, pp. 928-940, 2020.

[13] B. Wilie, K. Vincentio, G. I. Winata, S. Cahyawijaya, X. Li, Z. Y. Lim, S. Soleman, R. Mahendra, P. Fung, S. Bahar and A. Purwarianti, "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," 2019.

[14] C. A. Putri, "Analisis Sentimen Review Film Berbahasa Inggris Dengan Pendekatan Bidirectional Encoder Representations from Transformers," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 6, no. 2, pp. 181-193, 2020.

[15] G. B. Herwanto, A. Maulida Ningtyas, K. E. Nugraha and I. Nyoman Prayana Trisna, "Hate Speech and Abusive Language Classification using fastText," *2019 2nd International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2019*, no. December 2019, pp. 69-72, 2019.

## 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.

- [16] M. Lewis, "A Cognitive Linguistics Overview of Offense and Hate Speech," *SSRN Electronic Journal*, 2013.
- [17] S. T. B and J. L. Mitranont, *Automatic Discovery of Abusive Thai Language Usages in Social Networks* Suppawong, 2017.
- [18] A. F. Hidayatullah, A. A. Fadila, K. P. Juwairi and R. A. Nayoan, "Identifikasi Konten Kasar Pada Tweet Bahasa Indonesia," *Jurnal Linguistik Komputasional*, vol. 2, no. 1, pp. 1-5, 2019.
- [19] C. Nobata, J. Tetreault, A. Thomas, Y. Mehdad and Y. Chang, "Abusive language detection in online user content," *25th International World Wide Web Conference, WWW 2016*, pp. 145-153, 2016.
- [20] S. Russell and P. Norvig, *AI a modern approach*, vol. 2, 2010, p. 1151.
- [21] I. Goodfellow Yoshua Bengio, A. Courville, I. Goodfellow and Y. Bengio, *5 Deep Learning Title: Deep learning*, 2016, p. 777.
- [22] Y. Y. Chen, Y. H. Lin, C. C. Kung, M. H. Chung and I. H. Yen, "Design and implementation of cloud analytics-assisted smart power meters considering advanced artificial intelligence as edge analytics in demand-side management for smart homes," *Sensors (Switzerland)*, 2019.
- [23] Y. Lecun, Y. Bengio and G. Hinton, "Deep learning," *Nature*, 2015.
- [24] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin, "Attention is all you need," *Advances in Neural Information Processing Systems*, 2017.
- [25] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, vol. 17, no. 1, pp. 168-192, 2018.
- [26] D. J. M. Pasaribu, K. Kusrini and S. Sudarmawan, "Peningkatan Akurasi Klasifikasi Sentimen Ulasan Makanan Amazon dengan Bidirectional LSTM



**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.

dan Bert Embedding," *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, vol. 10, no. 1, pp. 9-20, 2020.

- [27] I. Abu Farha and W. Magdy, "Multitask Learning for {A}rabic Offensive Language and Hate-Speech Detection," *Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools, with a Shared Task on Offensive Language Detection*, no. May, pp. 86-90, 2020.
- [28] S. T. Luu, H. P. Nguyen, K. van Nguyen and N. L. T. Nguyen, "Comparison between traditional machine learning models and neural network models for Vietnamese hate speech detection," *arXiv*, 2020.
- [29] P. Badjatiya, S. Gupta, M. Gupta and V. Varma, "Deep learning for hate speech detection in tweets," *26th International World Wide Web Conference 2017, WWW 2017 Companion*, pp. 759-760, 2017.
- [30] D. P. N. Lyrawati, "DETEKSI UJARAN KEBENCIAN PADA TWITTER MENJELANG PILPRES 2019 DENGAN MACHINE," *Jurnal Ilmiah Matematika Volume 7 No.3*, 2019.
- [31] I. Alfina, R. Mulia, M. I. Fanany and Y. Ekanata, "Hate speech detection in the Indonesian language: A dataset and preliminary study," *2017 International Conference on Advanced Computer Science and Information Systems, ICACISIS 2017*, Vols. 2018-January, pp. 233-237, 5 2018.
- [32] T. T. A. Putri, "Analisis dan deteksi hate speech pada sosial twitter berbahasa indonesia," *Master's thesis, Faculty of Computer Science*, no. Universitas Indonesia, 2018.
- [33] M. O. Ibrohim and I. Budi, "A Dataset and Preliminaries Study for Abusive Language Detection in Indonesian Social Media," *Procedia Computer Science*, vol. 135, pp. 222-229, 1 2018.

## LAMPIRAN A

Berikut ini merupakan dataset beserta label yang digunakan pada penelitian

No	Tweet	HS	Abusive	HS_Individual	HS_Group	HS_Religion	HS_Race	HS_Physical	HS_Gender	HS_Other	HS_Weak	HS_Moderate	HS_Strong
1	- disaat semua cowok berusaha melacak perhatian gue. loe lantas remehkan perhatian yg gue kasih khusus ke elo. basic elo cowok bego !!!'	1	1	1	0	0	0	0	0	1	1	0	0
2	RT USER: USER siapa yang telat ngasih tau elu?edan sarap gue bergaul dengan cigax jifla calis sama siapa noh licew juga'	0	1	0	0	0	0	0	0	0	0	0	0
3	41. Kadang aku berfikir, kenapa aku tetap percaya pada Tuhan padahal aku selalu jatuh berkali-kali. Kadang aku merasa Tuhan itu ninggalkan aku sendirian. Ketika orangtuaku berencana berpisah, ketika kakakku lebih memilih jadi Kristen. Ketika aku anak ter	0	0	0	0	0	0	0	0	0	0	0	0
4	USER USER AKU ITU AKU\n\nKU TAU MATAMU SIPIT TAPI DILIAT DARI MANA ITU AKU'	0	0	0	0	0	0	0	0	0	0	0	0
5	USER USER Kaum cebong kapor udah keliatan dongoknya dari awal tambah dongok lagi hahahah'	1	1	0	1	1	0	0	0	0	0	1	0
6	USER Ya bani taplak dkk \xf0\x9f\x98\x84\xf0\x9f\x98\x84 \xf0\x9f\x98\x84'	1	1	0	1	0	0	0	0	1	0	1	0
7	deklarasi pilkada 2018 aman dan anti hoax warga dukuh sari jabon	0	0	0	0	0	0	0	0	0	0	0	0
8	Gue baru aja kelar re-watch Aldnoah Zero!!! paling kampret emang endingnya! 2 karakter utama cowonya kena friendzone bray! XD URL	0	1	0	0	0	0	0	0	0	0	0	0

### 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. 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.

9	Nah admin belanja satu lagi port terbaik nak makan Ais Kepal Milo, Ais Kepal Horlicks atau Cendol Topping kaw kaw. ❖??❖ Doket mano tu ? Gerai Rojak Mertuaku - Taipan 2 (depan TWINS BABY ROMANTIKA / Bank Islam Senawang) ❖???	0	0	0	0	0	0	0	0	0	0	0	0
10	USER Enak lg klo smbil ngewe'	0	1	0	0	0	0	0	0	0	0	0	0
11	Setidaknya gw punya jari tengah buat lu, sebelum gw ukur nyali sama bacot lu \xf0\x9f\x98\x8f'	1	1	1	0	0	0	0	0	1	1	0	0
12	USER USER USER USER BANCI KALENG MALU GA BISA JAWAB PERTANYAAN KAMI DARI 2 HARI LALU.... NYUNGSEP KOE USER URL	1	1	1	0	0	0	0	1	0	1	0	0
13	Kalo belajar ekonomi mestinya jago memprivatisasi hati orang. Duh.. ironi USER	0	0	0	0	0	0	0	0	0	0	0	0
14	Aktor huruhara 98 Prabowo S ingin lengserkan pemerintahan Jokowi.... Nyata	1	0	1	0	0	0	0	0	1	1	0	0
15	USER Bu guru enakan jadi jablay atau guru esde sih.\nKayaknya menikmati jadi pecun ini guru.'	1	1	1	0	0	0	0	1	0	1	0	0
16	USER USER USER USER USER USER Lawan bicara gw gak intelek kyk loe, yg otak gak punya. Ttg kencing onta, gw mengakui hadis nabi dan itu sahih. Kafir kyk loe pasti menolak, makanya loe ahlunnar.'	1	1	1	0	1	0	0	0	0	1	0	0
17	Belakangan ini kok fikiran ampas banget ya'	0	0	0	0	0	0	0	0	0	0	0	0
18	Ari sarua beki mah repeh monyet\xf0\x9f\x98\x86\xf0\x9f\x98\x86'	0	1	0	0	0	0	0	0	0	0	0	0
19	Jadi cowo itu harus Gantle kalo ga Gantle itu namanya BANCI !'	1	1	0	1	0	0	0	1	0	0	1	0
20	USER Slga mnr bom \xf0\x9f\x98\x82'	0	0	0	0	0	0	0	0	0	0	0	0
21	Asw ya tapi gua jarang ngambek, tacut wkkwkwkwkw gua kan bucin.'	0	1	0	0	0	0	0	0	0	0	0	0
22	USER kalo kamu noob pasti peluang disakitin nya lebih gede si'	0	0	0	0	0	0	0	0	0	0	0	0



**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.

23	USER Joko Widodo dinilai sebagai presiden terlemah dalam sejarah Indonesia. Hal ini terjadi bukan saja karena Jokowi tidak memiliki modal dukungan politik yang cukup, melainkan juga karena ketidakmampuan pemerintahannya menghadapi situasi ekonomi g	0	0	0	0	0	0	0	0	0	0	0	
24	PELAJAR SMA KEC BILAH HILIR DEKLARASI ANTI HOAX PILKADA 2018 DAMAI	0	0	0	0	0	0	0	0	0	0	0	
25	Bandara Udara Internasional Kertajati dibangun oleh Gubernur Ahmad Heryawan tanpa gunakan APBN pemerintah pusat. Insya Allah siap beroperasi bulan depan	0	0	0	0	0	0	0	0	0	0	0	
26	Siapapun gubernur dan presidennya, rakyatnya, ya kita <sup>2</sup> juga...	0	0	0	0	0	0	0	0	0	0	0	
27	Ini si USER kerjanya delay mulu! Setan!	1	1	0	1	0	0	0	0	1	0	1	0
28	menurutku pintu sorga ada yaitu pintu sorga yang asli dan pintu hatimu modusbanget	0	0	0	0	0	0	0	0	0	0	0	0
29	RT USER USER USER PKI hanya muncul jika jelang pemilihan kepala daerah dan pemilihan presiden krn digoreng sampai gosong.	1	0	0	1	0	1	0	0	0	0	1	0
30	.....												
31 69	USER Mana situ ngasih("): itu cuma foto ya kutil onta'	1	1	1	0	0	0	0	0	1	1	0	0

## LAMPIRAN B

Berikut ini merupakan hasil prediksi label tweet serta perbandingan sama atau tidak dengan label tweet yang ada di dataset pada penelitian ini :

Keterangan : Hasil Label yang di tebakkan merupakan label yang berbeda dengan label data test.

No	Tweet	HS	Abusive	HS_Individual	HS_Group	HS_Religion	HS_Race	HS_Physical	HS_Gender	HS_Other	HS_Weak	HS_Moderate	HS_Strong
1	Om...Ahoknya...Mana Om,Katanya Si BABi Ahok Bisa Ngatasi Banjir...Om...BABi Jgn Membual Om...?!?!	1	1	1	0	0	0	0	0	1	1	0	0
2	USER Sini aku jilatin sayang memek mu'	0	1	0	0	0	0	0	0	0	0	0	0
3	Me every night:\nAAAAA ANJENG GUE MALU BANGSAD. MAU ILANG AJA. AAAAAA. APAAN SIH GUE JIJK BANGET. AAAAAA. INI APAAN SIH GUE DULU NGGAK DANTA. AAAAAA. MAU ILANG AJA. AAAAAA.'	0	1	0	0	0	0	0	0	0	0	0	0
4	#NowPepatahPetitih bersama Sucipto Hadi Purnomo - dosen budaya Jawa di Jurusan Sastra Jawa Unnes. Tema ĩ½??Aja Gumunan, Aja Getunan, Aja Kagetan, Aja Alemani;½??.	0	0	0	0	0	0	0	0	0	0	0	0
5	USER \xe2\xa0\x80\n\xe2\xa0\x80\xe2\xa0\x80\xe2\xa0\x80\xe2\xa0\x80\xe2\xa0\x80 \xe3\x80\x8c Saya! Saya! Yang dulu bisu, kini sudah bisa bagi cuap-cuap! \xe3\x80\x8d'	0	0	0	0	0	0	0	0	0	0	0	0
6	METRO TV sedang sibuk acara belain JOKOWI TIDAK OTORITER...faktanya RAKYAT YG MERASAKAN.. kebijakan yg main gunakan PERPPU semaunya saja..	1	0	1	0	0	0	0	0	1	1	0	0

### 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. 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.

7	Bodo ah, males buat pacaran. Gue yg jomblo napa lu yg repot.'	0	0	0	0	0	0	0	0	0	0	0	0
8	USER Lg bisnis biarin aja yg pting 2019ganti presiden yg amanah seswe janji..gmgnya gk impor malah serba impor.petani negri sendiri gigit jari. ampun deh'	1	0	1	0	0	0	0	0	1	1	0	0
9	Tetap fokus bersama Kang USER menangkan pasangan #Asyik USER USER menjadi Gubernur Jabar. Abaikan yg lain! #AsyikM3nang #AherUntukIndonesia	0	0	0	0	0	0	0	0	0	0	0	0
10	Berdayakan Ekonomi Pesantren Ulama Jawa Barat Dukung Jokowi di Pilpres 2019 #TetapJokowi2019	0	0	0	0	0	0	0	0	0	0	0	0
11	RT USER: Goblok loe . itu org yg berbeda Cong ! URL	1	1	1	0	0	0	0	0	1	1	0	0
12	USER Hutang karena belanja lebih besar dari pendapatan, DPR RI menyetujui APBN. DPR RI HRUS mengontrol pengeluaran ( efisien, efektif) jgn boros, perangi korupsi, bantu pemerintah (baca rakyat Indonesia) utk naikan pend.pajak.'	0	0	0	0	0	0	0	0	0	0	0	0
13	Partai Aceh Target 50 Kursi Parlemen Hingga Memenangkan Pilkada 2018 -	0	0	0	0	0	0	0	0	0	0	0	0
14	RT USER: Mainstream sdh jelas,bhw oposisi mengesankan negeri ini sdg "Bermasalah"\n\nDr mulai "horor"utang, kemiskinan, nge-fans sm Puti\Xe2\x80\xa6'	0	0	0	0	0	0	0	0	0	0	0	0
15	USER nanya seriusan mbah... Kenapa meng "iya" kan maen film di Kafir ini? Ceritanya menarik? Sesuai dengan hati nurani?'	0	0	0	0	0	0	0	0	0	0	0	0
16	USER Gue kira lu tidur, ternyata lu mengetik tai onta banyak banyak y :---)'	1	1	1	0	0	0	0	0	1	1	0	0
17	USER USER Etnis cina memang ada yg ikut berjuang tapi lebih sangat banyak yg jadi penghianat jd antek belanda..	1	0	0	1	0	1	0	0	0	0	1	0
18	Ketika ngomong, Sylvi keliatan oke. Tapi begitu didebat Ahok, Sylvi langsung kekuatan goblok.	1	1	1	0	0	0	0	0	1	1	0	0

19	USER USER Salah sendiri milih gubernur yg diusung dr isu sara	0	0	0	0	0	0	0	0	0	0	0	0
20	kenapa anak babi jalan nya nunduk???karena malu liat ulah bapak nya yang ngebelain ahok..	1	1	1	0	0	0	0	0	1	1	0	0
21	"Tidak peduli sesulit apa aku menghadapi sesuatu, aku akan tetap tersenyum seperti orang idiot" - Chanyeol EXO K'	0	0	0	0	0	0	0	0	0	0	0	0
22	RT USER: #RETWEET \nGadis Bandung Cantik Ngentot\nSama Pacar Di Hotel..\n\n#horny #memek #colmek #colokmemek #jembut #memekbecek #n\xe2\x80\xa6	0	1	0	0	0	0	0	0	0	0	0	0
23	USER USER Trus hub nya putusan MK dan Presiden apa?? Mengingat kekuasaan di Republik ini sdh dipisah menjadi 3 yaitu eksekutif, legislatif dan yudikatif dmn ketiganya berdiri sendiri..	0	0	0	0	0	0	0	0	0	0	0	0
24	USER Tetap aja SALAH dimata bani kampret.. katanya tinggal tunggu kontrak abis, jadi msh ada peluang nyolong klo JKW ga 2 periode, diperpanjang lg kontraknya, makin tebellah kantong2 tikus berdasar'	1	1	0	1	0	0	0	0	1	0	1	0
25	USER Lagi tunangan ama Ak di cina'	0	0	0	0	0	0	0	0	0	0	0	0
26	Setelah nonton Disobedience aku jadi penasaran kenapa wanita Yahudi Orthodox pada pake wig'	0	0	0	0	0	0	0	0	0	0	0	0
27	USER A. 2013 tepatnya tanggal 22 september 2013\n#AFFU19Indosiar #timnasDay'	0	0	0	0	0	0	0	0	0	0	0	0
28	USER Pasti g dukung, aye diem2 itu mengagumi duet jokowi dan puan untuk jd capres cawapres loh, suer! #2019GantiPresiden	1	0	1	0	0	0	0	0	1	1	0	0
29	Apalagi teriak Allah hu Barbar, rusak telinga Bung Fahri, kapan diskusi Buang Islam di DPR?	1	1	0	1	1	0	0	0	0	0	1	0
...	.....												
31	Pria Belanda Ini Jadi Mualaf Usai Ikuti Perjalanan Islam Nusantara Oleh NU di Eropa	0	0	0	0	0	0	0	0	0	0	0	0

**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 RIWAYAT HIDUP

### INFORMASI PERSONAL



Nama	Reski Saputra
NIM	11651101881
Jurusan	Teknik Informatika
Agama	Islam
E-mail	11651101881@students.uin- suska.ac.id

### INFORMASI PENDIDIKAN

Tahun 2002-2009	SDN 052 Pekanbaru
Tahun 2009-2012	SMPN 25 Pekanbaru
Tahun 2012-2015	SMAN 2 Siak Hulu
Tahun 2016-2021	UIN Sultan Syarif Kasim Riau S1 Teknik Informatika

#### 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.