

**KLASIFIKASI *CYBERBULLYING* PADA MEDIA SOSIAL KOMENTAR
INSTAGRAM BAHASA INDONESIA MENGGUNAKAN METODE
*BIDIRECTIONAL LONG-SHORT TERM MEMORY (BILSTM)***

SKRIPSI

SRI WAHYUNI

191402045



**PROGRAM STUDI S1 TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA
MEDAN
2025**

**KLASIFIKASI *CYBERBULLYING* PADA MEDIA SOSIAL KOMENTAR
INSTAGRAM BAHASA INDONESIA MENGGUNAKAN METODE
*BIDIRECTIONAL LONG-SHORT TERM MEMORY (BILSTM)***

SKRIPSI

Diajukan untuk melengkapi tugas dan memenuhi syarat memperoleh ijazah
Sarjana Teknologi Informasi

SRI WAHYUNI

191402045



**PROGRAM STUDI S1 TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA**

MEDAN

2025

PERSETUJUAN

Judul : KLASIFIKASI CYBERBULLYING PADA MEDIA
SOSIAL KOMENTAR INSTAGRAM BAHASA
INDONESIA MENGGUNAKAN METODE
BIDIRECTIONAL LONG-SHORT TERM MEMORY
(BILSTM)

Kategori : SKRIPSI

Nama : SRI WAHYUNI

Nomor Induk Mahasiswa : 191402045

Program Studi : SARJANA (S1) TEKNOLOGI INFORMASI

Fakultas : ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA

Medan, 6 Januari 2025

Komisi Pembimbing:

Pembimbing 2

Pembimbing 1

Fahrurrozi Lubis B.IT., M.Sc.IT

NIP. 198610122018052001

Mohammad Fadly Syah Putra M.Sc

NIP. 198301292009121003

Diketahui/disetujui oleh

Program Studi S1 Teknologi Informasi

Ketua,

Dedy Arisandi S.T., M.Kom.

NIP. 197908312009121002

PERNYATAAN

KLASIFIKASI *CYBERBULLYING* PADA MEDIA SOSIAL KOMENTAR INSTAGRAM
BAHASA INDONESIA MENGGUNAKAN METODE *BIDIRECTIONAL LONG-SHORT*
TERM MEMORY (BILSTM)

SKRIPSI

Saya mengakui bahwa skripsi ini merupakan hasil karya Saya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing telah disebutkan sumbernya.

Medan, 6 Januari 2025

SRI WAHYUNI

191402045

UCAPAN TERIMA KASIH

Alhamdulillahirobbil'alamin, dengan mengucapkan syukur kepada Allah Subhanahu wa Ta'ala atas segala nikmat dan karunia-Nya, penulis dapat menyelesaikan skripsi ini sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Teknologi Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara. Adapun judul skripsi ini adalah "*Klasifikasi Cyberbullying pada Komentar Instagram Berbahasa Indonesia Menggunakan Metode Bidirectional Long-Short Term Memory (BiLSTM)*".

Penulis juga ingin menyampaikan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan doa, dukungan, dan bantuan selama proses penulisan skripsi ini. Pada kesempatan ini, saya ingin mengungkapkan rasa syukur dan terima kasih yang mendalam kepada:

1. Kedua orang tua penulis, Ayahanda Sahrul dan Ibunda Sri Milakawati yang telah membesarkan penulis serta selalu memberikan doa yang terbaik dan dukungan kepada penulis dari awal Pendidikan hingga menyelesaikan skripsi ini.
2. Semua keluarga penulis yang telah memberikan berbagai motivasi, semangat serta doa kepada penulis sehingga penulis dapat menyelesaikan skripsi ini.
3. Bapak Prof. Dr. Muryanto Amin S.Sos, M.Si. yang merupakan Rektor Universitas Sumatera Utara.
4. Ibu Dr. Maya Silvi Lydia, B.Sc., M.Sc. selaku Dekan Fakultas Ilmu Komputer dan Teknologi Informasi.
5. Bapak Mohammad Fadly Syah Putra M.Sc selaku Dosen Pembimbing 1 dan Bapak Fahrurrozi Lubis B.IT., M.Sc.IT selaku Dosen Pembimbing 2 yang telah membimbing penulis serta memberikan kritik, saran, dan motivasi dalam penulisan skripsi ini.
6. Bapak Dedy Arisandi S.T., M.Kom. selaku Ketua Program Studi S1 Teknologi Informasi, Universitas Sumatera Utara.
7. Bapak Ivan Jaya S.Si., M.Kom selaku Sekretaris Program Studi S1 Teknologi Informasi, Universitas Sumatera Utara.
8. Seluruh Dosen Program Studi S1 Teknologi Informasi yang telah memberikan banyak ilmu yang sangat bermanfaat bagi penulis selama masa perkuliahan.
9. Seluruh pegawai dan staff Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara yang telah membantu penulis dalam segala urusan administrasi dan akademik selama masa perkuliahan penulis.

10. Sahabat serta teman seperjuangan penulis selama masa perkuliahan Katherin Anna Patherisia Lesnussa, Nanda Amelia, Tritia Mutiara dan Putri Laura Latersia yang telah banyak membantu, menyemangati, penghibur, pendengar keluh kesah penulis selama masa perkuliahan sampai penyelesaian skripsi penulis.
11. Teman-teman dari Kom C serta teman-teman seangkatan dari stambuk 2019 Teknologi Informasi Universitas Sumatera Utara yang telah berjuang selama masa perkuliahan.
12. Seluruh senior dan junior yang sudah menyemangati penulis selama masa perkuliahan.
13. Pihak lainnya yang sudah membantu penulis dalam menyelesaikan skripsi ini yang tidak dapat dijabarkan semuanya.

Penulis mengucapkan terimakasih kepada Tuhan yang maha esa atas semua orang yang telah mendukung, menemani, mendoakan, membimbing serta memberikan wawasan kepada penulis selama proses penyusunan skripsi. Semoga Tuhan Yang Maha Esa dapat membalas kebaikan kepada semua pihak yang membantu penulis menyelesaikan skripsinya.

Medan, 6 Januari 2025

Sri Wahyuni

191402045

ABSTRAK

Salah satu negara terbanyak yang menggunakan media sosial di dunia salah satunya negara Indonesia. Berdasarkan data dari *katadata.co.id*, pengguna media sosial di Indonesia mencapai 96 juta pada tahun 2017 dan diperkirakan meningkat lebih dari 25,3 juta pada tahun 2022, mencerminkan peningkatan lebih dari 30%. Survei yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) mengungkapkan bahwa sekitar 49% pengguna internet menjadi korban *cyberbullying* di media sosial. Selain itu, Komisi Perlindungan Anak Indonesia (KPAI) menerima ratusan laporan tentang *cyberbullying* setiap tahun. Penelitian ini bertujuan untuk menemukan kata atau kombinasi kata yang memiliki potensi terbesar untuk melakukan *cyberbullying* di media sosial, khususnya Instagram. Dataset yang digunakan dalam penelitian ini terdiri dari 8000 komentar berbahasa Indonesia yang dikumpulkan melalui aplikasi Instagram menggunakan *library instaloader*. Metode yang digunakan untuk klasifikasi adalah algoritma *Bidirectional Long-Short Term Memory* (BiLSTM). Hasil penelitian menunjukkan bahwa algoritma *BiLSTM* mampu mengklasifikasikan komentar *cyberbullying* dengan tingkat akurasi sebesar 95.5%. Model ini membagi data latih dan data uji dengan proporsi 80% dan 20%, dan mampu mengklasifikasikan komentar ke dalam empat kategori: *sexism*, *flaming*, *body shaming*, dan netral. Berdasarkan hasil tersebut, model BiLSTM menunjukkan performa yang baik dan dapat diandalkan untuk klasifikasi komentar *cyberbullying* di media sosial Instagram berbahasa Indonesia.

Kata kunci : *Cyberbullying*, Instagram, *Bidirectional Long-Short Term Memory* (BiLSTM), Klasifikasi Komentar, Bahasa Indonesia

***CLASSIFICATION OF CYBERBULLYING ON SOCIAL MEDIA INSTAGRAM
COMMENTS IN INDONESIAN USING THE BIDIRECTIONAL LONG SHORT TERM
MEMORY (BiLSTM) METHOD***

ABSTRACT

Indonesia is one of the countries with the highest number of social media users in the world. According to data from katadata.co.id, the number of social media users in Indonesia reached 96 million in 2017 and is estimated to increase by more than 25.3 million by 2022, reflecting an increase of over 30%. A survey conducted by the Indonesian Internet Service Providers Association (APJII) revealed that approximately 49% of internet users have been victims of cyberbullying on social media. Additionally, the Indonesian Child Protection Commission (KPAI) receives hundreds of reports of cyberbullying each year. A GlobalWebIndex study showed that 61% of respondents use social media during their leisure time. This study aims to identify words or combinations of words with the highest potential for cyberbullying on social media, particularly Instagram. The dataset used in this study consists of 8000 Indonesian-language comments collected through the Instagram application using the `instaloader` library. The classification method employed is the Bidirectional Long Short Term Memory (BiLSTM) algorithm. The results indicate that the BiLSTM algorithm can classify cyberbullying comments with an accuracy rate of 95.5%. The model divides the training and testing data with a proportion of 80% and 20%, respectively, and can classify comments into four categories: sexism, flaming, body shaming, and neutral. Based on these results, the BiLSTM model demonstrates good performance and reliability for classifying cyberbullying comments on Indonesian-language Instagram.

Keywords : Cyberbullying, Instagram, Bidirectional Long-Short Term Memory (BiLSTM), Comment Classification, Indonesian Language

DAFTAR ISI

PERSETUJUAN	i
PERNYATAAN	ii
UCAPAN TERIMA KASIH.....	iii
ABSTRAK	v
<i>ABSTRACT</i>	vi
DAFTAR ISI.....	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Batasan Masalah.....	4
1.5 Manfaat Penelitian.....	4
1.6 Metodologi Penelitian	4
1.7 Sistematika Penulisan.....	6
BAB 2 LANDASAN TEORI	7
2.1 <i>Bullying</i>	7
2.2 <i>Cyberbullying</i> ,	8
2.3 <i>Natural Language Processing (NLP)</i>	13
2.4 <i>Text Preprocessing</i>	14
2.5 <i>Text Classification</i>	14
2.6 <i>Recurrent Neural Networks (RNN)</i>	15
2.7 <i>Long Short-Term Memory (LSTM)</i>	16
2.8 <i>Bidirectional Long Short-Term Memory (BiLSTM)</i>	21
2.9 <i>Word Embedding</i>	23
2.10 Penelitian Terdahulu.....	25

BAB 3 ANALISIS DAN PERANCANGAN SISTEM	30
3.1 Dataset.....	30
3.2 Arsitektur Umum.....	31
3.2.1 <i>Input Dataset</i>	32
3.2.2 <i>Pre-processing</i>	33
3.2.3 <i>Word Embedding</i>	36
3.2.4 Pembangunan Model	38
3.2.5 Melatih Model.....	40
3.3 Perancangan Sistem.....	41
3.3.1 Rancangan Tampilan Bagian Beranda (<i>Home</i>).....	41
3.3.2 Rancangan Tampilan Bagian <i>Training</i>	42
3.3.3 Rancangan Tampilan Bagian <i>Testing</i>	42
3.3.4 Rancangan Tampilan Bagian <i>User Input</i>	43
3.3.5 Rancangan Tampilan Bagian Input Link	44
3.4 Metode Evaluasi.....	45
BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM.....	46
4.1 Implementasi Sistem	46
4.1.1 Spesifikasi Perangkat yang Digunakan.....	46
4.1.2 Implementasi Rancangan Tampilan Antarmuka	47
4.2 Implementasi Model.....	52
4.2.1 Model Training.....	52
4.2.2 Model Testing	54
4.3 Metode Evaluasi.....	55
BAB 5 KESIMPULAN DAN SARAN.....	64
5.1 Kesimpulan.....	64
5.2 Saran.....	64
DAFTAR PUSTAKA	65

DAFTAR TABEL

Tabel 2.1 <i>Keyword Flaming</i>	10
Tabel 2.2 <i>Keyword Seksisme</i>	11
Tabel 2.3 <i>Keyword Body Shaming</i>	12
Tabel 2.4 <i>Penelitian Terdahulu</i>	26
Tabel 3.1 <i>Hasil Crawling Komentar Instagram</i>	30
Tabel 3.2 <i>Komentar Yang Dilebeli</i>	32
Tabel 3.3 <i>Penerapan Cese Folding</i>	33
Tabel 3.4 <i>Penerapan Cleaning</i>	34
Tabel 3.5 <i>Penerapan Normalization</i>	34
Tabel 3.6 <i>Penerapan Stopword Removal</i>	35
Tabel 3.7 <i>Penerapan Steaming</i>	35
Tabel 3.8 <i>Penerapan Tokenization</i>	36
Tabel 3.9 <i>Confusion Matrix</i>	45
Tabel 4.1 <i>Percobaan Pengubahan Hyperparameter</i>	53
Tabel 4.2 <i>Hasil Pengujian Sistem</i>	54
Tabel 4.3 <i>Class Dataset</i>	55
Tabel 4.4 <i>Keterangan Confusion Matrix Class Sexism</i>	57
Tabel 4.5 <i>Keterangan Confusion Matrix Class Flaming</i>	58
Tabel 4.6 <i>Keterangan Confusion Matrix Class Body Shaming</i>	59
Tabel 4.7 <i>Keterangan Confusion Matrix Class Netral</i>	60
Tabel 4.8 <i>Perhitungan Kalkulasi</i>	61

DAFTAR GAMBAR

Gambar 2.1 <i>Contoh Komentar Flaming</i>	9
Gambar 2.2 <i>Contoh Komentar Seksisme</i>	11
Gambar 2.3 <i>Contoh Komentar Body Shaming</i>	12
Gambar 2.4 <i>Alur Pengolahan Bahan Alami</i>	13
Gambar 2.5 <i>Model Klasifikasi Teks</i>	15
Gambar 2.6 <i>Jaringan Neural Berulang</i>	16
Gambar 2.7 <i>Jaringan Neural Berulang Yang Terbuka Gulungannya</i>	16
Gambar 2.8 <i>Modul Berulang RNN Yang Berisi Satu Lapisan</i>	16
Gambar 2.9 <i>Modul Berulang LSTM Berisi Empat Layer</i>	17
Gambar 2.10 <i>Cell State Jaringan LSTM</i>	17
Gambar 2.11 <i>Gates pada LSTM</i>	17
Gambar 2.12 <i>Layer Sigmoid</i>	18
Gambar 2.13 <i>Gate Layer & Tanh Layer</i>	19
Gambar 2.14 <i>Pembuatan Cell State Baru</i>	20
Gambar 2.15 <i>Menentukan Hasil Output</i>	20
Gambar 2.16 <i>Arsitektur BiLSTM</i>	22
Gambar 2.17 <i>Model BiLSTM</i>	22
Gambar 2.18 <i>One-hot Encoding kata dalam kosa kata</i>	24
Gambar 3.1 <i>Arsitektur Umum</i>	31
Gambar 3.2 <i>Representasi Vector Berdasarkan Kata “bodoh”</i>	37
Gambar 3.3 <i>Penggalan Code Embedding Layer</i>	38
Gambar 3.4 <i>Pemenggalan Code Bidirectional LSTM Layer</i>	39
Gambar 3.5 <i>Pemenggalan Code Dropout Layer</i>	39
Gambar 3.6 <i>Pemenggalan Code Dense Layer</i>	40
Gambar 3.7 <i>Penggalan Code Learned Model</i>	40
Gambar 3.8 <i>Home Page (Rancangan)</i>	41
Gambar 3.9 <i>Training Page (Rancangan)</i>	42
Gambar 3.10 <i>Testing Page (Rancangan)</i>	43

Gambar 3.11 <i>User Input Page (Rancangan)</i>	44
Gambar 3.12 <i>Input Link Page (Rancangan)</i>	44
Gambar 4.1 <i>Home Page (Tampilan)</i>	47
Gambar 4.2 <i>Training Page (Tampilan)</i>	48
Gambar 4.3 <i>Training Page (Tampilan Hasil Preprocessing)</i>	48
Gambar 4.4 <i>Training Page (Tampilan Grafik Loss dan Accuracy)</i>	48
Gambar 4.5 <i>Testing Page (Tampilan)</i>	49
Gambar 4.6 <i>Tampilan Halaman Akurasi dan Confussion Matrix (Testing Page)</i>	49
Gambar 4.7 <i>Testing Page (Tampilan Hasil Klasifikasi)</i>	50
Gambar 4.8 <i>User Input Page (Tampilan)</i>	50
Gambar 4.9 <i>User Input Page (Tampilan Hasil Inputan Teks)</i>	51
Gambar 4.10 <i>Input Link Page (Tampilan)</i>	51
Gambar 4.11 <i>Input Link (Tampilan Hasil dari Inputan Link)</i>	52
Gambar 4.12 <i>Grafik Loss dan Accuracy</i>	53
Gambar 4.13 <i>Confusion Matrix (Semua Class)</i>	56
Gambar 4.14 <i>Konversi Matriks ke Class 0 (Sexism)</i>	57
Gambar 4.15 <i>Konversi Matriks ke Class 1 (Flaming)</i>	58
Gambar 4.16 <i>Konversi Matriks ke Class 2 (Body Shaming)</i>	59
Gambar 4.17 <i>Konversi Matriks ke Class 3 (Netral)</i>	60

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Salah satu negara dengan jumlah penggunaan media sosial tertinggi di dunia adalah Indonesia. Menurut data dari *katadata.co.id*, pada tahun 2017, jumlah penggunaan media sosial di Indonesia mencapai angka 96 juta. Di tahun 2022, diperkirakan jumlah penggunaan media sosial kian meningkat menjadi 25,3 juta, menunjukkan peningkatan melebihi dari angka 30%.

Data dari *Napoleon Cat* mengungkapkan bahwa pada Desember 2022, pengguna Instagram di Indonesia mencapai 97,17 juta, meningkat sebesar 0,10% dari bulan sebelumnya yang berjumlah 97,07 juta pengguna. Pengguna Instagram di Indonesia didominasi oleh wanita dengan proporsi 51,5%, sementara 48,5% adalah laki-laki. Dari segi usia, 39% pengguna Instagram yang berusia 18 dan 24 tahun, 28,8% berusia 25 dan 34 tahun, 11,5% berusia 13-17 tahun, dan 13,2% berada dalam rentang usia 35-44 tahun.

Menurut *katadata.co.id*, survei yang dilaksanakan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) menunjukkan hasil sekitar 49% penggunaan internet di Indonesia telah menjadi korban *cyberbullying* dalam media sosial. Selain itu, Komisi Perlindungan Anak Indonesia (KPAI) melaporkan bahwa mereka menerima ratusan laporan mengenai *cyberbullying*. Dengan populasi mencapai 150 juta orang, penggunaan media sosial di Indonesia sangat bervariasi. Studi dari *GlobalWebIndex* yang dilaporkan oleh Lidwina pada tahun 2019 menunjukkan bahwa 61% responden menggunakan media sosial di waktu luang mereka.

Menurut Willard (2007), *cyberbullying* bisa diklasifikasikan dalam beberapa jenis, yaitu *Flaming*, yang merujuk pada tindakan mengirim atau memposting pesan ofensif melalui internet, mencakup bahasa yang kasar, vulgar, menghina, dan terkadang mengancam. Pelecehan (*Harrassment*) adalah perilaku ofensif terhadap individu yang biasanya dikirim melalui saluran komunikasi pribadi seperti email, pesan instan, dan pesan teks. Pelecehan ini umumnya dianggap tidak pantas menurut standar sosial dan moral karena merendahkan, menghina, atau mempermalukan seseorang. Pencemaran nama baik (*Denigration*) adalah penyebaran informasi berbahaya, tidak benar, atau kejam mengenai seseorang, dengan tujuan merusak pertemanan atau reputasi target. Rasis (*Racism*) merupakan bentuk kepercayaan yang

menganggap bahwasannya ras ataupun budaya tertentu lebih unggul atau bahkan memandang rendah ras atau budaya lainnya. Rasisme dapat mencakup terhadap stereotip yang negative, perlakuan yang tidak seimbang, sikap yang berupaya untuk saling memisahkan diri, ataupun mencakup kekerasan. Seksisme (*sexism*) yang merupakan bentuk dari prasangka ataupun diskriminasi terhadap suatu gender. Seksisme ini sering kali terjadi dari peran gender yang tradisional dan stereotip. Fenomena ini mencakup berbagai keyakinan ataupun sikap yang beranggapan bahwasannya suatu jenis kelamin memiliki nilai yang lebih dibandingkan dengan yang lainnya, hal ini biasanya disebut dalam kasus misogini (kebencian terhadap perempuan) ataupun misandri (kebencian terhadap laki-laki), yang memberikan rasa ketidakpercayaan terhadap lawan jenis. Selain itu, terdapat pula penghinaan fisik (*body shaming*), yang merupakan tindakan mengejek atau merendahkan penampilan fisik dari seseorang. Ruang lingkup *body shaming* ini sangatlah luas mulai dari ejekan terhadap berat badan (baik itu gemuk maupun kurus), tinggi badan, kondisi rambut, bentuk tubuh, penampilan wajah, serta beberapa kondisi medis yang meninggalkan bekas fisik seperti psoriasis.

Penelitian oleh *Microsoft Research* pada bulan Mei-April 2020 menggunakan "*Digital Civility Index*" (Indeks Peradaban Digital) menunjukkan tingkat kesopanan pengguna internet di dunia digital. Hasilnya, netizen Indonesia berada di peringkat ke-29 dari bawah, atau ketiga terbawah menjadi "netizen tidak sopan". Analisis yang dilaksanakan di 32 negara dengan 16.000 responden, dan ada 503 pengguna internet dari Indonesia. Tambahan informasi dari laporan Polda Metro Jaya menunjukkan bahwa setidaknya 25 kasus dilaporkan setiap hari, dengan jumlah yang terus bertambah seiring dengan meningkatnya pengguna internet dan penanganan kasus cyberbullying yang kurang efektif.

Penelitian mengenai mengklasifikasikan komentar *bullying* di Instagram dengan menggunakan *k-nearest neighbor* (KNN) dengan menggunakan 1000 data yang mana datanya dibagi menjadi 2 yaitu 500 data untuk data *bullying* dan 500 data lagi untuk data *non-bullying*. Hasil dari penelitian ini menunjukkan nilai akurasi dengan rata-rata mencapai angka 58,83% (Candra dan rozana, 2020).

Penelitian lainnya tentang klasifikasi komentar *cyberbullying* pada Instagram yang menggunakan metode *Lexicon-Based* dan *Naïve Bayes Classifier* (studi kasus: pemilihan presiden Indonesia tahun 2019) dengan menggunakan 360 data, dengan pembagian 80:20 antara data pelatihan dan data pengujian, datanya dibagi menjadi 179 data untuk *cyberbullying* dan ada 181 data untuk *non-cyberbullying*. Hasil yang dihasilkan dalam penelitian ini mendapatkan nilai akurasi yang mencapai 97% (Syarif et al., 2019).

Penelitian oleh Fadli dan Hidayatullah (2021) tentang mengidentifikasi *cyberbullying* di media sosial Twitter menggunakan metode LSTM dan BiLSTM, mengumpulkan 6.835 data yang menunjukkan kecenderungan antara *cyberbullying* dan *non-cyberbullying*. Kedua algoritma *deep learning* yang diterapkan, LSTM dan BiLSTM, menunjukkan kinerja yang relatif sama. Keakuratan masing-masing algoritma adalah 93,77% untuk *Long-Short Term Memory* dan 95% untuk *Bidirectional Long-Short Term Memory*. *F1-Score* untuk *Long-Short Term Memory* adalah 92,02% dan untuk *Bidirectional Long-Short Term Memory* adalah 93,84%.

Berdasarkan hasil penelitian tersebut, penulis melakukan pendekatan guna menemukan metode terbaik dalam mengidentifikasi komentar *cyberbullying* di media sosial Instagram. Maka dari itu, penulis membuat penelitian dengan judul “*Klasifikasi Cyberbullying Pada Komentar Media Sosial Instagram Bahasa Indonesia Menggunakan Metode Bidirectional Long-Short Term Memory (BiLSTM)*”. Diharapkan hasil penelitian ini dapat menciptakan sistem yang mampu mengidentifikasi komentar *cyberbullying* dengan akurasi tinggi berdasarkan isi komentar.

1.2 Rumusan Masalah

Cyberbullying telah menjadi perhatian utama masyarakat karena banyak kasusnya yang sering dikaitkan dengan bunuh diri. *Cyberbullying* dapat terjadi dalam berbagai bentuk seperti rasisme, seksisme, penghinaan fisik, dan pencemaran nama baik. *Cyberbullying Research Center*, sebuah lembaga *non-profit*, menyatakan bahwa bentuk *bullying* yang dilakukan remaja, secara langsung ataupun daring, bisa menyebabkan trauma, tindakan bunuh diri, dan percobaan pembunuhan. Mengingat dampak yang berbahaya dari *cyberbullying*, maka diperlukan langkah-langkah pencegahan dan deteksi dini melalui klasifikasi komentar di Instagram untuk menentukan apakah komentar tersebut termasuk *cyberbullying* atau tidak, sehingga korban dan pelaku dapat dilindungi.

1.3 Tujuan Penelitian

Mengidentifikasi kata atau kombinasi kata yang memiliki potensi terbesar untuk digunakan dalam *cyberbullying* di media sosial, khususnya Instagram. Penemuan ini akan mempermudah proses klasifikasi komentar *cyberbullying* dalam bahasa Indonesia di Instagram dengan menerapkan algoritma *Bidirectional Long-Short Term Memory* (BiLSTM) yang merupakan tujuan utama dari penelitian ini.

1.4 Batasan Masalah

Disini penulis membatasi ruang lingkup masalah agar tidak terlalu luas. Berikut adalah beberapa batasan dari hal-hal tersebut, yaitu:

1. Dataset yang digunakan diambil dengan cara mengcrawling data dari platform media sosial Instagram.
2. Komentar yang diambil menggunakan bahasa Indonesia.
3. Sistem hanya memproses komentar yang dimasukkan ke dalam file CSV (*Comma Seperated Value*).
4. Hasil klasifikasi aspek terbatas pada empat aspek yaitu: aspek *flaming*, seksisme, *body shaming* dan netral serta hasil klasifikasi komentar terbatas pada *cyberbullying* dan *non-cyberbullying*.
5. Dalam melakukan klasifikasi sistem hanya menggunakan algoritma *Bidirectional Long-Short Term Memory (BiLSTM)*.
6. Evaluasi hanya berfokus pada akurasi dan bukan kecepatan.

1.5 Manfaat Penelitian

Dalam penelitian ini memiliki beberapa manfaat, antara lain:

1. Memberikan wawasan kepada pengguna media sosial Instagram tentang komentar yang tergolong dalam *cyberbullying*.
2. Memungkinkan pengguna, orang tua, kerabat, dan pemerintah untuk meminimalkan perbuatan *cyberbullying*.
3. Tersedianya informasi tentang kosa serta gabungan kata-kata yang berpotensi besar terjadinya *bullying* di media sosial khususnya Instagram.
4. Untuk mengetahui kinerja metode *Bidirectional Long-Short Term Memory (BiLSTM)* untuk mengklasifikasikan aspek dan mengidentifikasi komentar *cyberbullying* bahasa Indonesia dari Instagram.

1.6 Metodologi Penelitian

Uraian metode penelitian ini menjelaskan tentang panduan atau tahapan dalam melakukan penelitian. Tujuan penerapan metodologi ini adalah untuk memperoleh hasil sesuai dengan yang diharapkan. Berikut merupakan tahapan penelitian yang dilakukan:

1. Tinjauan Pustaka

Langkah pertama dalam penelitian ini adalah mengumpulkan, membaca dan mencatat penelitian pustaka ataupun informasi Pustaka dan mengelola bahan-bahan yang berkaitan dengan penelitian ini. Referensi berasal dari publikasi ilmiah, artikel, buku, serta video tentang pembelajaran mesin, pengolahan teks, klasifikasi berbasis aspek, *cyberbullying*, dan penggunaan algoritma pada Pemrosesan Bahasa Alami (*Natural Language Processing*), *fastText*, dan metode *Bidirectional Long-Short Term Memory (BiLSTM)*.

2. Analisis Masalah

Tahap kedua adalah menganalisis masalah yang bertujuan untuk membuat dasar bagi proses perencanaan dan penentuan prioritas selanjutnya. Setelah analisis masalah dilakukan dan rumusan masalah ditemukan, pada tahap ini akan dihubungkannya metode *Bidirectional Long-Short Term Memory (BiLSTM)* sebagai solusi dalam mengurangi masalah tersebut.

3. Pengumpulan Data

Data yang akan diperlukan dalam penelitian ini dikumpulkan dan akan dipastikan bahwasannya kualitas data yang telah dikumpulkan sudah sesuai dengan tujuan penelitian.

4. Perancangan Sistem

Tahap ketiga dari penelitian ini adalah perancangan sistem, dimana sistem dirancang berdasarkan hasil kajian Pustaka dan analisis permasalahan yang ada.

5. Pengimplementasian Sistem

Tahap implementasi algoritma adalah langkah keempat akan dilakukan untuk menyelesaikan rancangan sistem yang telah disepakati. Dalam penelitian ini, algoritma *BiLSTM* diterapkan sebagai solusi untuk mengatasi permasalahan yang ada.

6. Pengujian Sistem

Tahap kelima dalam metode penelitian ini adalah melakukan pengujian sistem untuk mengetahui performa sistemnya dalam melakukan pengujian sistem terhadap data yang belum dikenal dengan menerapkan algoritma *Bidirectional Long-Short Term Memory (BiLSTM)* sebelumnya.

7. Pembuatan Laporan

Pada tahapan ini akan dilakukannya penyusunan laporan yang komprehensif mengenai seluruh proses penelitian hingga mencapai kesimpulan akhir. Setelah penelitian ini telah selesai dilakukan maka laporan ini akan menyajikan gambaran lengkap sesuai dengan proses yang telah dilakukan.

1.7 Sistematika Penulisan

Berikut ini adalah struktur sistematika yang digunakan dalam penyusunan laporan ini, yang dijelaskan secara terperinci seperti berikut ini:

Bab 1 : Pendahuluan

Di dalam bab pertama ini akan membahas alasan dilakukannya penelitian ini, masalah yang akan diselesaikan, manfaat dan tujuan dilaksanakannya penelitian, batasan di dalam penelitian ini, penggunaan metodologi dan sistematika atau struktur penulisan laporan.

Bab 2 : Landasan Teori

Bab ini mencakup tinjauan literatur tentang teori-teori yang mendukung penelitian ini, penjelasan tentang algoritma dan prapemrosesan yang digunakan, dan penelitian sebelumnya akan menjadi validitas dalam penyusunan skripsi ini.

Bab 3 : Analisis dan Perancangan

Pada bab ini, pembahasan mencakup penggunaan dataset, perancangan metode *Bidirectional Long-Short Term Memory*, dan desain antarmuka sistem.

Bab 4 : Implementasi dan Pengujian

Bab ini membahas penerapan hasil perancangan dari bab sebelumnya, serta pengujian yang dilakukan untuk memastikan keefektifan model yang dikembangkan. Selain itu, hasil dan analisis pengujian sistem *Bidirectional Long-Short Term Memory* juga akan di bahas.

Bab 5 : Kesimpulan dan Saran

Bab ini menyajikan ringkasan dari penelitian yang telah dilaksanakan beserta hasil-hasilnya, serta memberikan rekomendasi atau masukan yang berguna untuk penyempurnaan penelitian lebih lanjut.

BAB 2

LANDASAN TEORI

2.1 *Bullying*

Menurut *Kemenpppa* (2016), berbagai bentuk intimidasi atau eksploitasi yang dilakukan secara terencana atau tidak oleh satu individu atau kelompok individu yang merasa bahwa dirinya lebih kuat atau berkuasa terhadap individu lainnya dengan maksud untuk menyakiti dan dilakukan secara berulang disebut *bullying*, yang dikenal juga sebagai “penindasan/risak”.

Bullying merupakan bentuk perilaku kekerasan yang dilakukan oleh individu atau kelompok yang lebih “kuat” kepada individu atau kelompok yang dianggap “lemah” akan diganggu secara fisik maupun secara psikologis. Pelaku *bullying*, yang biasanya disebut sebagai *bully*, merasa memiliki otoritas dalam melakukan apa pun atas korbannya. Dan korban pun akan merasa terancam oleh pelaku dan merasa tidak berdaya (Djuwita, 2005; Ariesto, 2009).

Bullying dapat berdampak untuk semua orang yang terlibat mulai dari anak yang menjadi korban, yang merupakan pelaku, maupun orang yang menyaksikan *bullying* tersebut terjadi. Dampak negatif ini dapat mempengaruhi keadaan fisik dan mental anak-anak. Untuk gambaran yang lebih ekstrem, *bullying* bisa memicu tindakan yang sangat fatal seperti bunuh diri. Adapun dampak dari *bullying* sebagaimana diuraikan oleh *Kemenpppa* (2016) adalah:

a Konsekuensi untuk victim (korban *bullying*)

Korban *bullying* dapat mengalami perubahan sikap maupun perilaku. Salah satu dampak signifikan dari tindakan *bullying* terhadap korban adalah masalah kesehatan mental seperti merasakan kecemasan yang berlebihan, depresi, dan *Post-Traumatic Stress Disorder* (PTSD). Korban yang mengalami dampak kesehatan mental ini akan menghadapi periode pemulihan yang panjang.

b Konsekuensi untuk bullies (pelaku *bullying*)

Pelaku *bullying* cenderung merasakan kepercayaan diri dan memiliki harga diri yang tinggi, serta tidak toleran terhadap frustrasi yang dialami oleh korban. Mereka memiliki keinginan lebih untuk mengendalikan orang lain dan kurangnya rasa empati terhadap korbannya. Dengan melakukan *bullying* akan memberikan kesan kepada pelaku bahwa mereka memiliki kontrol atas situasi. Jika pelaku didiamkan terus untuk melakukan

bullying, mereka dapat berkembang menjadi individu yang melakukan kekerasan atau tindakan kejahatan lainnya.

c Konsekuensi untuk menyaksikan tindakan *bullying* (*bystanders*)

Tindakan *bullying* yang diabaikan tanpa adanya tindakan, maka untuk individu lain yang menyaksikan akan menganggap bahwa *bullying* merupakan perilaku yang dapat diterima dalam masyarakat. Akibatnya, sejumlah orang mungkin akan ikut bergabung dalam tindakan *bullying* karena merasa takut akan menjadi sasaran, atau yang lebih parah, mereka mungkin diam saja tanpa mengambil tindakan karena merasa tidak perlu menghentikannya.

2.2 Cyberbullying

Cyberbullying merujuk pada semua tindakan kekerasan yang dilakukan oleh anak atau remaja terhadap teman sebayanya melalui jejaring sosial atau dunia maya. Ketika seorang mengalami penghinaan, intimidasi, bahkan dipermalukan oleh anak lain melalui media sosial, teknologi digital, atau telepon seluler maka itu juga bisa disebut tindakan dari *cyberbullying* (Endah, 2016).

Cyberbullying dapat berwujud dalam beragam bentuk dan cara. Misalnya, mengirim ancaman melalui pesan teks, menyebarkan foto yang memermalukan korban, membuat situs web dengan tujuan menyebarkan berita bohong dan mengejek korban, membobol akun media sosial seseorang untuk mengancam korban, hingga mengakibatkan kerusakan. Motif pelaku *cyberbullying* sangat beragam, mulai dari rasa marah, dendam, frustrasi, mencari perhatian, hingga sekadar mengisi waktu luang. Seringkali, motif pelaku hanyalah untuk bersenang-senang.

Cyberbullying lebih mudah untuk dilakukan dibandingkan dengan kekerasan yang konvensional, yang mana pelaku tidak perlu berhadapan secara langsung dengan korbannya untuk melancarkan aksinya. Pelaku yang berada di balik layar komputer maupun ponsel tidak akan melihat secara langsung dampak yang diberikannya terhadap korban, sehingga membuat pelaku dengan senang mengatakan berbagai hal yang bertujuan untuk merendahkan dan mengintimidasi korbannya. Selain itu, sulit bagi anak-anak atau remaja untuk mengenali *cyberbullying*, contohnya orang tua atau guru, karena mereka sering menggunakan kode-kode yang terdiri dari akronim atau emotikon yang tidak dapat dimengerti oleh orang lain. Perlu diingat bahwa kasus *cyberbullying* sering kali tidak terdeteksi. Korban sering enggan mengakui

bahwa mereka menjadi korban *bullying* karena mereka khawatir akses internet dan ponsel mereka akan dibatasi jika melaporkan kejadian tersebut.

Karena beragamnya tindakan yang dikategorikan sebagai *cyberbullying*, seringkali seseorang tidak menyadari bahwa mereka melakukan atau menjadi korban *cyberbullying*. Ketika seseorang menjadi korban *bullying*, mereka biasanya akan dianggap sebagai sasaran. *Bullying* dapat menyebabkan trauma, kecemasan, depresi, dan bahkan kematian. *Cyberbullying* dapat menyerang seseorang tanpa memandang agama atau status sosialnya. Peningkatan depresi dapat disebabkan oleh tekanan atau intimidasi, baik secara fisik maupun verbal, namun riset yang dilakukan menunjukkan korban yang mengalami *cyberbullying* memiliki tingkat depresi yang lebih besar dibandingkan dengan korban yang mengalami kekerasan secara verbal (Liu et al., 2020).

Menurut *Pew Research Center* (2017), satu dari sepuluh orang menjadi sasaran *bullying* dikarenakan penampilan fisik mereka (9%), ras (8%), atau jenis kelamin (8%). *Bullying* juga dilakukan karena alasan lain, seperti agama (5%) atau orientasi seksual (3%). Bentuk *cyberbullying* yang sering terjadi di media sosial meliputi:

- *Flaming*

Flaming adalah istilah yang mengacu pada hubungan yang bermusuhan, marah, serta menghina yang seringkali berupa dalam serangan pribadi yang menyakitkan. *Flaming* dapat terjadi di banyak tempat, seperti forum online, papan diskusi, ruang obrolan, email dan pesan instan. Ketika orang marah, mereka sering menggunakan huruf kapital, seperti “KAMU IDIOT DAN AKU BENCI KAMU!”. Banyak pesan *flaming* yang tidak hanya keji dan kejam tetapi juga sering mengabaikan fakta atau alasannya melakukan *flaming* pada orang tersebut (Bauman, 2014).



Gambar 2.1 Contoh Komentar *Flaming*

Mengutip dari (Chamidah & Sahawaly, 2021) Tabel 2.1 menunjukkan *keyword* yang dapat digunakan untuk mengklasifikasikan komentar flaming.

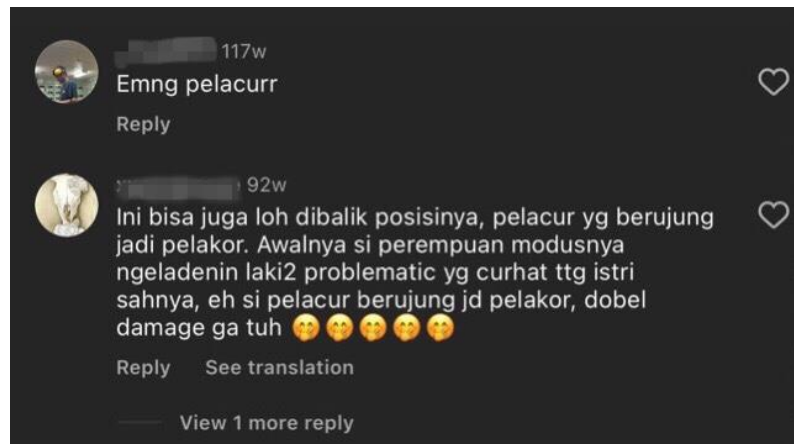
Tabel 2.1 *Keyword Flaming*

Kelas	Kata Kunci Bullying	
<i>Flaming</i>	<ul style="list-style-type: none"> - Anjing - Goblok - Bangsat - Orang - Anak 	<ul style="list-style-type: none"> - Bodoh - Tahu - Tai - Lihat - Baru

- Seksisme

Seksisme mencakup sikap, keyakinan, perilaku individu, praktik organisasi, dan budaya yang mencerminkan penilaian buruk terhadap individu berdasarkan gender atau menunjang ketidaksetaraan antara perempuan dan laki-laki. Secara esensial, seksisme diartikan sebagai pemikiran dan perilaku yang berlandaskan kriteria jenis kelamin, yang dapat membatasi tindakan kita sesuai dengan jenis kelamin tertentu (Thompson, 1995).

Menurut Ayres, Friedman & Leaper (2009) Komentar yang berkaitan dengan seksualitas sering terjadi pada Perempuan Dimana interaksi ini sering terjadi dalam sehari-hari. Hal ini menjadi salah satu bentuk pelecehan seksual atau objektif seksual. Tidak semua Perempuan melihat komentar seksisme dengan cara yang sama. Ini dapat terjadi dalam bentuk agresif, seperti “Pekerjaan berhitung tidak usah untuk dia, dia kan Perempuan, performa perempuan dalam hal matematika tidak bagus” atau bahkan dalam komentar yang bijak seperti “Biar saya saja yang bawa barang-barangnya, Perempuan tidak seharusnya membawa barang-barang yang berat” atau bahkan ada yang bersikap merendahkan seperti “Pakai baju yang ketat saja, kamu akan terlihat seksi jika mengenakannya”. Karena tidak semua perempuan menanggapi komentar ini dengan cara yang sama, komentar semacam itu kadang-kadang muncul dalam situasi yang ambigu, memungkinkan perempuan untuk menganggapnya sebagai pujian atau bahkan penghinaan.



Gambar 2.2 Contoh Komentar Seksisme

Melalui beberapa tinjauan literatur, beberapa kata kunci yang berkaitan dengan seksisme (diskriminasi gender) sering digunakan. Tabel 2.2 keyword yang paling sering muncul untuk menggambarkan seksisme.

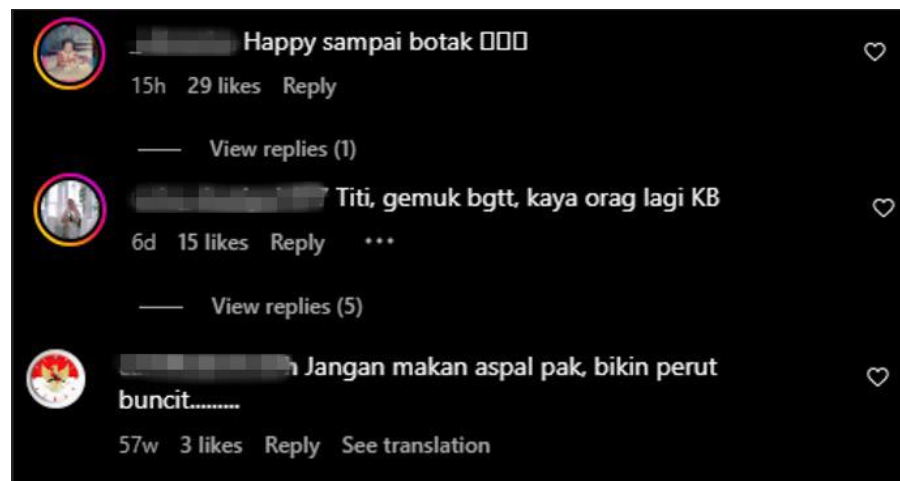
Tabel 2.2 Keyword Seksisme

Keyword	Cara Penggunaan	Referensi
Jalang	Perempuan terus dipermalukan dengan label yang bertentangan dengan standar normative ini, yang dilestarikan oleh pelaku secara sengaja atau tidak sengaja.	Felmlee, Rodis & Zhang (2019).
Pelacur		
Vagina		
Penggoda	Penghinaan penuh kebencian	Johnson (2014).
Malapetaka		
Jalang Gemuk		
Payudara	Objektifikasi Perempuan dan pelecehan	Ciampaglia (2017).
Seperti Perempuan	Meminimalkan prestasi perempuan	Thaliakr (2016).
Tomboy		
Murah		

- *Body Shaming*

Menurut Pane (2023) mengomentari dan menghina penampilan fisik orang lain dapat disebut sebagai *body shaming*, yang mana hal itu termasuk kedalam *bullying*. Mereka yang melakukan *body shaming* mungkin melakukannya karena berbagai alasan, seperti hanya ingin membuat suasana menjadi lebih lucu, ingin menjadikan bahan bercandaan, atau memang dengan niat ingin menghina.

Masa remaja yang mana termasuk kedalam fase dimana seseorang menjadi semakin sadar akan bagaimana mereka terlihat. Dengan munculnya gagasan “*influencer*” di media sosial dan banyaknya pemberitaan tentang bagaimana seharusnya berpenampilan, remaja yang mudah terpengaruh dan belum dewasa semakin terjebak dalam konsep tentang “tubuh ideal”. Hal ini lah yang membuat orang berspekulasi bahwa orang yang tidak memiliki tubuh ideal akan dipermalukan (Eisenberg et al, 2003).



Gambar 2.3 Contoh Komentar *Body Shaming*

Mengutip dari Chamidah & Sahawaly (2021) Tabel 2.3 menunjukkan *keyword* yang dapat digunakan untuk mengklasifikasikan komentar *body shaming*.

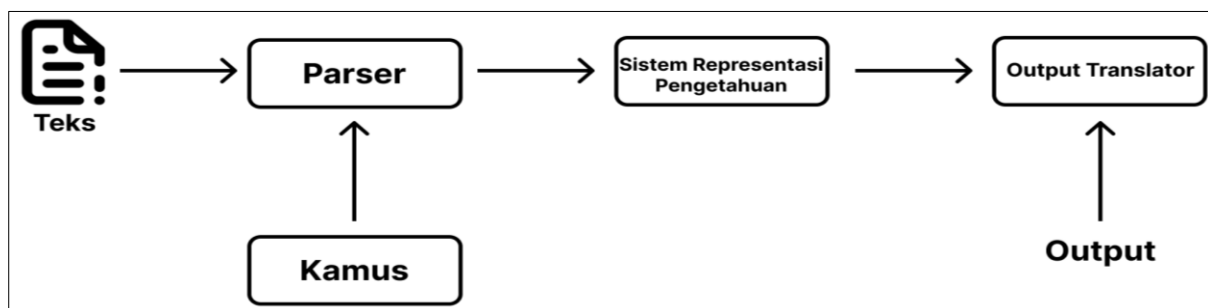
Tabel 2.3 *Keyword Body Shaming*

Kelas	Kata Kunci Bullying	
Body Shaming	- Cebol	- Cungkring
	- Jelek	- Dekil
	- Botak	- Sipit
	- Gendut	- Buncit
	- Item	- Orang

2.3 *Natural Language Processing (NLP)*

Rumaisa et al. (2021) menyatakan bahwa mengajarkan perangkat agar memahami, memproses, serta menghasilkan bahasa merupakan bagian dari *NLP*, yang merupakan cabang dari kecerdasan buatan (AI). Teknologi ini mendukung mesin pencari, layanan terjemahan otomatis, dan asisten suara. Saat ini, *Natural Language Processing* menjadi bagian penting dari kehidupan sehari-hari, karena ditemukan dalam asisten virtual (*siri*, *Google Assistant*, dan *Alexa*). *NLP* juga memiliki peranan yang penting dalam dunia industri bisnis yang berguna dalam memperoleh keunggulan dalam persaingan teknologi. Aplikasi yang menerapkan *NLP* memiliki kemampuan untuk mempermudah dalam berbagai aspek kehidupan, apalagi dalam menganalisis dan menghasilkan nilai dari data yang tidak terstruktur.

Natural Language Processing digunakan dalam sistem untuk memahami, menginterpretasikan dan menghasilkan bahasa manusia, sehingga memungkinkan adanya interaksi secara alami dan efektif antara manusia dan mesin. *NLP* bekerja dengan sistem yang akan menuliskan serta mendokumentasikan kaidah manusia, lalu akan disusun menjadi kalimat yang formal dan terstruktur, sehingga sistem dapat dengan mudah memanipulasi dan memberikan respons yang tepat. Berikut ini adalah komponen utama NLP menurut Soyusiawaty dan Haspiyan (2015):



Gambar 2.4 Alur Pengolahan Bahan Alami
(Soyusiawaty dan Haspiyan, 2015)

a. *Parser*

Parser merupakan serangkaian perintah, yang biasanya berupa program, dibagi menjadi bagian yang lebih mudah diproses, diuji sintaksnya, dan kemudian dilampirkan ke tag yang menunjukkan masing-masing bagian. Parser berguna untuk mendapatkan teks dan memecahnya menjadi bagian-bagian gramatikal seperti subjek, verba, nomina, dan objek.

b. *Sistem Representasi Pengetahuan*

Sistem representasi pengetahuan merupakan komposisi untuk mengatur cara mengkodekan dan menyimpan pengetahuan yang diperoleh dalam basis pengetahuan. Sistem ini digunakan untuk menganalisis keluaran suatu parser dan menentukan maknanya, dalam hal ini sistem yang dimaksud adalah kamus yang berisikan kata-kata bahasa alami.

c. *Output Translator*

Translator akan memberikan teks yang seimbang dalam bahasa target, baik berupa teks maupun audio, dengan begitu memungkinkan komunikasi lintas bahasa yang lebih efektif.

2.4 *Text Preprocessing*

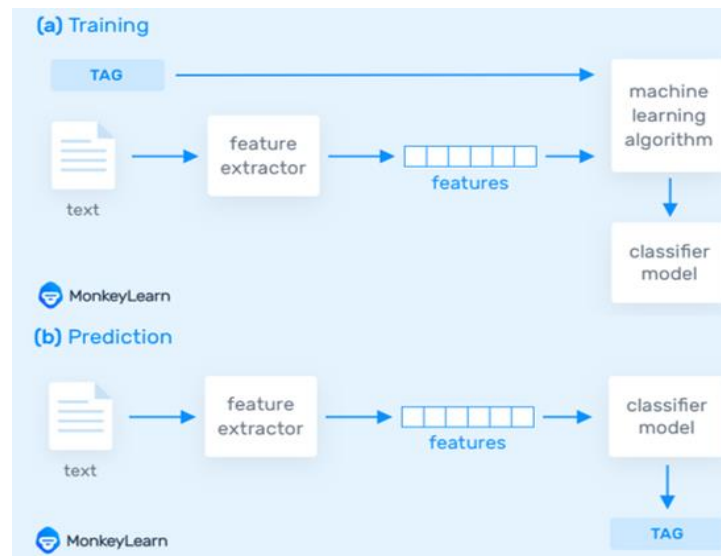
Dalam *TextMining*, *Natural Language Processing* (NLP), dan *Information Retrieval* (IR) tahapan *preprocessing* merupakan tahapan yang cukup penting. Dalam *TextMining*, mengekstraksi kebijaksanaan yang menarik dan penting dari data teks yang tidak terstruktur akan dilakukan dalam proses *preprocessing*. Untuk memenuhi kebutuhan informasi pengguna, *informasi retrieval* (IR) menentukan dokumen mana yang harus diambil dari koleksi (Kannan, 2014). Akibatnya, pengambilan keputusan akan menciptakan perbandingan query dengan indeks (kata atau frasa) yang ditemukan dalam dokumennya.

Pada langkah ini, tindakan yang dilakukan ialah `toLowerCase`, membantu dan memastikan konsistensi serta keakuratan dalam pengelolaan *string* dengan mengonversi semua karakter huruf besar menjadi huruf kecil, dan tokenizing yang merupakan proses pemecahan deskripsi yang semula terdiri dari kalimat menjadi kata perkata, dan menghapus delimiter seperti titik (.), koma (,), spasi dan karakter angka yang terdapat dalam kata (Weiss et al, 2010).

2.5 *Text Classification*

Menurut *MonkeyLearn* (2023) teknik penting dalam penambangan teks adalah klasifikasi teks, yang melibatkan pembagian file teks mentah ke satu atau lebih kategori yang sudah ditentukan sebelumnya. Ini merupakan konsep langsung dari *machine learning*, yang mencakup deklarasi sekumpulan bilangan berlabel sebagai cara untuk menampilkan dokumen dan pengklasifikasian statistik yang dilatih dengan set pelatihan berlabel. Klasifikasi ini merupakan proses dimana sesuatu dikenali, dibedakan, dan dipahami. Proses ini menyebabkan

sesuatu dikelompokkan ke dalam kategori, biasanya dengan alasan tertentu. Prediksi, interferensi, dan pengambilan keputusan bergantung pada klasifikasi.



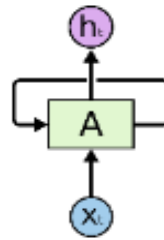
Gambar 2.5 Model Klasifikasi Teks
Sumber: Monkeylearn.com 2023

Model pembelajaran mesin dapat dimulai setelah data latih (*training*) dilatih dengan cukup memadai. Memperbaiki teks yang tidak terstruktur dengan rangkaian fitur yang terdapat dalam model klasifikasi untuk mendapatkan prediksi pada label, maka akan digunakan ekstraksi fitur yang sama.

Pembelajaran mesin dalam mengklasifikasikan teks umumnya lebih akurat dibandingkan dengan sistem buatan manusia, apalagi dalam tugas pengklasifikasian yang cukup kompleks. Sementara itu, klasifikasi dengan pembelajaran mesin lebih gampang untuk dipertahankan.

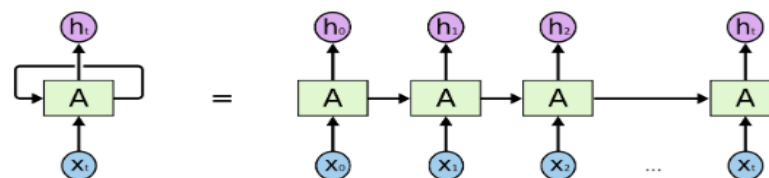
2.6 Recurrent Neural Networks (RNN)

Fauziyah et al (2022) Metode *deep learning* yang dapat melakukan pemrosesan data sekuensial secara berulang salah satunya adalah *Recurrent Neural Networks* (RNN). Dalam arsitektur *Recurrent Neural Networks* (RNN), banyak jaringan *neuron* berbalik dari dirinya sendiri atau ke *neuron* di *layer input*. Ini memungkinkan jaringan untuk menyimpan nilai, dan cara *input* menghasilkan nilai sebelumnya ke dalam jaringan dipengaruhi oleh arsitektur ini. Lapisan *input* mengimbangi lapisan pertama, dan lapisan berikutnya mengimbangi lapisan sebelumnya. Pada **Gambar 2.6** menunjukkan skema jaringan *neural* yang memiliki *loop*.



Gambar 2.6 Jaringan Neural Berulang
(Olah, 2015)

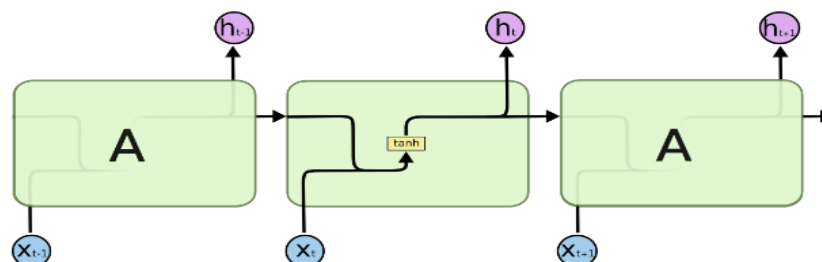
Dalam Gambar 2.6 di atas bagian jaringan saraf A menunjukkan nilai x_t menjadi inputan dan h_t adalah outputnya, dan terjadi perulangan (*loop*) dalam alurnya. *Recurrent Neural Networks* (RNN) dapat digambarkan sebagai *loop* dengan kemungkinan informasi berpindah dari satu jaringan ke jaringan berikutnya. Pada Gambar 2.7 menunjukkan hasil dari penjabaran *loop*.



Gambar 2.7 Jaringan Neural Berulang Yang Terbuka Gulungannya
(Olah, 2015)

2.7 Long Short-Term Memory (LSTM)

Komponen dari *RNN* yang memiliki kemampuan dalam mempelajari dependensi dalam jangka panjang salah satunya adalah *Long Short-Term Memory* (LSTM). Algoritma *LSTM* dipublikasikan oleh Hochreiter dan Schmidhuber ditahun 1997, dan kemudian disempurnakan dan dipopulerkan oleh banyak orang dalam karya berikutnya. Salah satu hal yang membedakan *LSTM* dan *RNN* adalah penambahan *memory cell*, yang merupakan alat yang sangat berguna dalam menyimpan informasi jangka waktu yang lama. Untuk pemrosesan data dengan *sequence* yang panjang, *LSTM* adalah alternatif yang dapat menangani masalah *vanishing gradient* pada *RNN*. Pada Gambar 2.8 arsitektur *RNN* menggunakan modul berulang yang memiliki konstruksi yang sangat sederhana, seperti lapisan *tanh* tunggal.



Gambar 2.8 Modul Berulang RNN Yang Berisi Satu Lapisan
(Olah, 2015)

Persamaan 2.1 menjelaskan persamaan *layer tanh*.

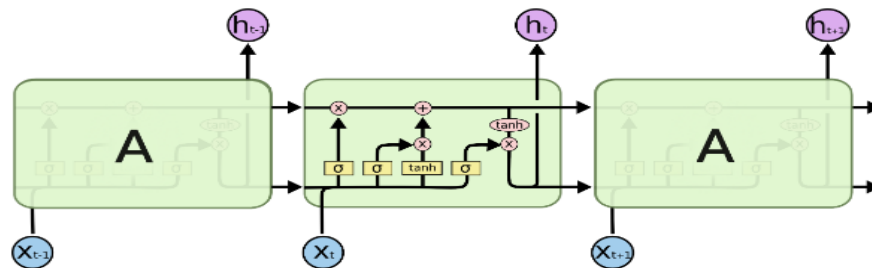
$$\text{Tanh}(x) = 2\sigma(2x) - 1 \quad (2.1)$$

Keterangan:

σ = Fungsi aktivasi sigmoid

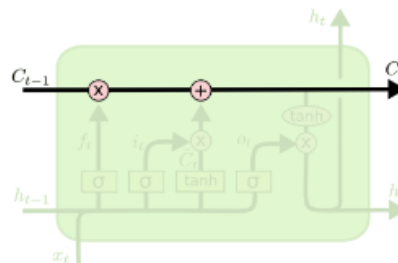
x = Data *input*

Sementara *LSTM* mempunyai struktur rantai yang sama, tetapi struktur *loop* yang dimiliki *LSTM* berbeda. Jaringan saraf yang dimiliki *LSTM* terdiri dari 4 *layer* yang berinteraksi dengan cara yang unik terdapat pada Gambar 2.9.



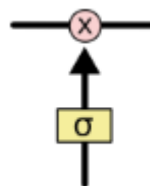
Gambar 2.9 Modul Berulang *LSTM* Berisi Empat Layer
(Olah, 2015)

Cell state yang merupakan bidang sejajar yang melewati bagian atas diagram dan menghubungkan semua *output layer* merupakan kunci dari *LSTM*, dapat dilihat Gambar 2.10.



Gambar 2.10 *Cell State* Jaringan *LSTM*
(Olah, 2015)

Menambahkan atau menghapus informasi ke keadaan *cell state*, dan keadaan *cell state* diatur secara terstruktur yang di sebut *gates* merupakan kemampuan yang dimiliki *LSTM*. *Sigmoid layer* dan *pointwise multiplication operation* yang merupakan bagian dari *Gates* terdiri dari yang ditunjukkan Gambar 2.11.



Gambar 2.11 *Gates* pada *LSTM*
(Olah, 2015)

Sigmoid layer akan mengeluarkan *output* dalam bentuk angka 0 dan 1, yang masing-masing akan menunjukkan beberapa informasi yang harus dibiarkan lewat. Nilai 0 menunjukkan bahwa “informasi tidak diteruskan”, sedangkan untuk menunjukkan bahwa “semua informasi akan diteruskan” akan diberi nilai 1. Persamaan *sigmoid layer* akan ditunjukkan pada persamaan 2.2.

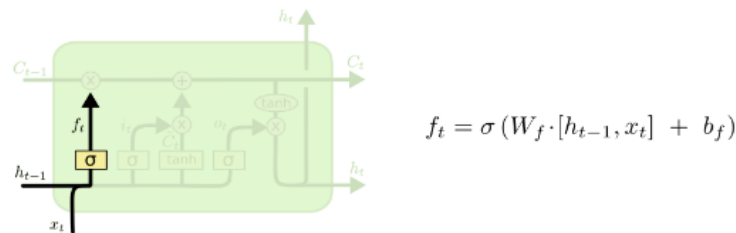
$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2.2)$$

Keterangan:

x = data *input*

e = konstanta (2,71828182845904)

Hal yang pertama kali dilakukan dalam *LSTM* adalah menentukan bahan mana yang akan dibuang dari *cell state*. *Layer sigmoid* yang dikenal sebagai “*forget gate layer*” menghasilkan keputusan ini. Pada kondisi *cell state* C_{t-1} , yang akan menangani h_{t-1} dan x_t sebagai *input* dan akan mengeluarkan *output* angka 0 atau 1 terjadi lapisan *forget gate*, dijelaskan Gambar 2.12.



Gambar 2.12 *Layer Sigmoid*
(Olah, 2015)

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.3)$$

Keterangan:

f_t = *Forget gate*

σ = Fungsi sigmoid

W_f = Nilai *weight* untuk *forget gate*

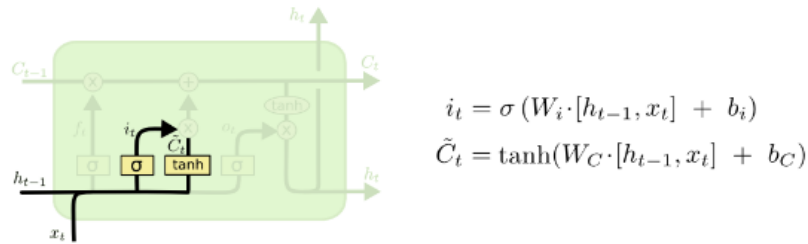
h_{t-1} = Nilai *output* sebelum orde ke t

x_t = Nilai *input* pada orde ke t

b_f = Nilai bias pada *forget gate*

Langkah berikutnya adalah menentukan bahan baru yang akan disimpan kedalam *cell state*. Tahapan ini terbagi menjadi dua bagian, *sigmoid layer* disebut “*input gate layer*” yang akan menentukan nilai mana yang akan diperbarui merupakan potongan pertama, dan “*tanh*

layer” akan menjadikan vektor nilai kandidat baru \tilde{C}_t yang ditambahkan ke dalam *cell state* merupakan bagian kedua. Gambar 2.13 menjelaskan proses berikut.



Gambar 2.13 Gate Layer & Tanh Layer
(Olah, 2015)

Persamaan *input gate* dijelaskan pada persamaan 2.4.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_f) \quad (2.4)$$

Keterangan:

i_t = *Input gate*

σ = *Fungsi sigmoid*

W_i = Nilai *weight* untuk *input gate*

h_{t-1} = Nilai *output* sebelum orde ke t

x_t = Nilai *input* pada orde ke t

b_f = Nilai bias pada *input gate*

Persamaan vektor nilai kandidat baru dijelaskan pada persamaan 2.5.

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (2.5)$$

Keterangan:

\tilde{C}_t = Nilai baru yang dapat ditambahkan ke *cell*

\tanh = fungsi *tanh*

W_C = Nilai *weight* untuk *cell state*

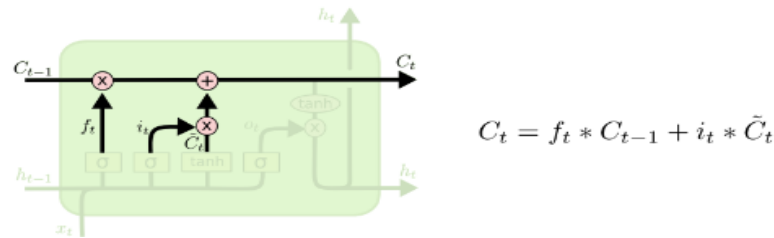
h_{t-1} = Nilai *output* sebelum orde ke t

x_t = Nilai *input* pada orde ke t

b_C = Nilai bias untuk *cell state*

Selanjutnya *cell state* akan diperbarui, C_{t-1} ke dalam *cell state* baru C_t . Dan akan melakukan penggandaan *cell state* yang lama dengan f_t , yang akan membuang bahan yang telah ditetapkan pada langkah *forget gate layer*. Setelah itu akan menambahkan $i_t * \tilde{C}_t$ ke dalam

nilai yang baru, yang dibandingkan berdasarkan berapa banyaknya keputusan untuk memulihkan setiap *state*. Gambar 2.14 penjabaran dari prosesnya.



Gambar 2.14 Pembuatan Cell State Baru
(Olah, 2015)

Persamaan *cell state* yang diperbarui dijabarkan dalam persamaan 2.6.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (2.6)$$

Keterangan:

C_t = cell state

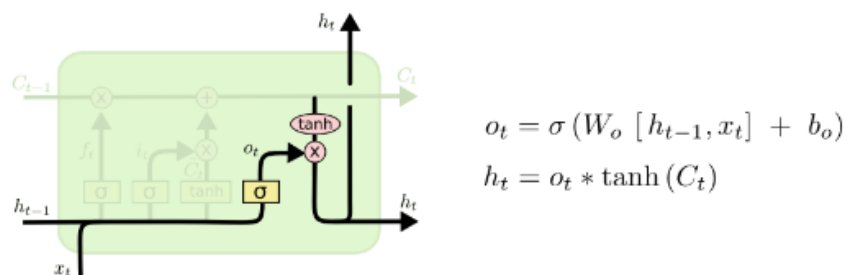
f_t = forget gate

C_{t-1} = cell state sebelum orde ke t

i_t = input gate

\tilde{C}_t = nilai baru yang ditambah ke cell state

Tahapan terakhir dari metode *LSTM* adalah menentukan hasil *output*. *Output* mana yang dikeluarkan akan berdasarkan *cell state*, tetapi dengan versi yang telah difilter. Ini dilakukan pada bagian awal *sigmoid layer*, yang menentukan potongan mana dari *cell state* yang akan dihasilkan. Setelah itu, diinput ke *tanh layer* guna mengganti nilai -1 dan 1 akan digandakan dengan *sigmoid gate*, sehingga hanya akan menampilkan *output* sesuai dengan sebelumnya. Tahapannya dijabarkan pada Gambar 2.15.



Gambar 2.15 Menentukan Hasil Output
(Olah, 2015)

Persamaan *output gate* akan dijabarkan pada persamaan 2.7.

$$o_t = \sigma (W_o [h_{t-1}, x_t] + b_o) \quad (2.7)$$

Keterangan:

o_t = *output gate*

σ = fungsi *sigmoid*

W_o = nilai *weight* untuk *output gate*

h_{t-1} = nilai *output* sebelum orde ke t

x_t = nilai *input* pada orde ke t

b_o = nilai bias pada *output gate*

Untuk persamaan nilai *output* pada orde ke t akan diuraikan pada persamaan 2.8.

$$h_t = o_t * \tanh (C_t) \quad (2.8)$$

Keterangan:

h_t = nilai *output* orde ke t

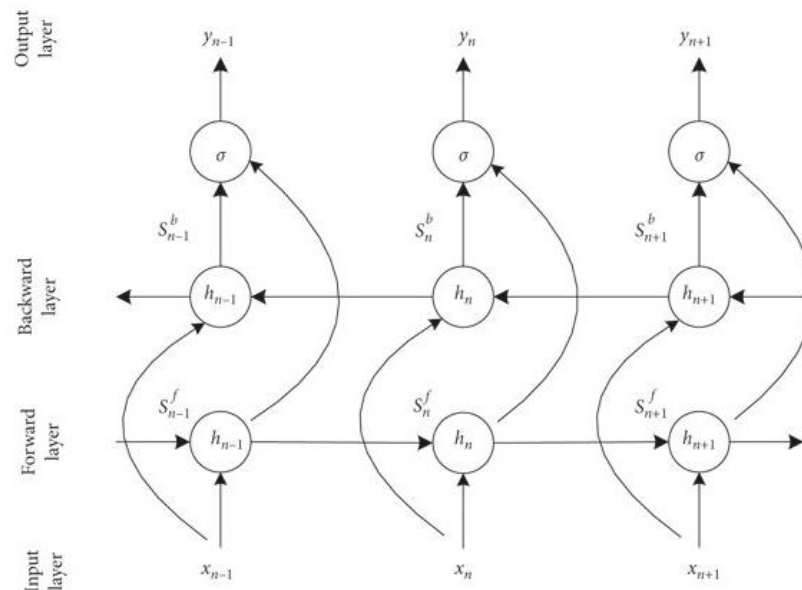
o_t = *output gate*

\tanh = fungsi *tanh*

C_t = *cell state*

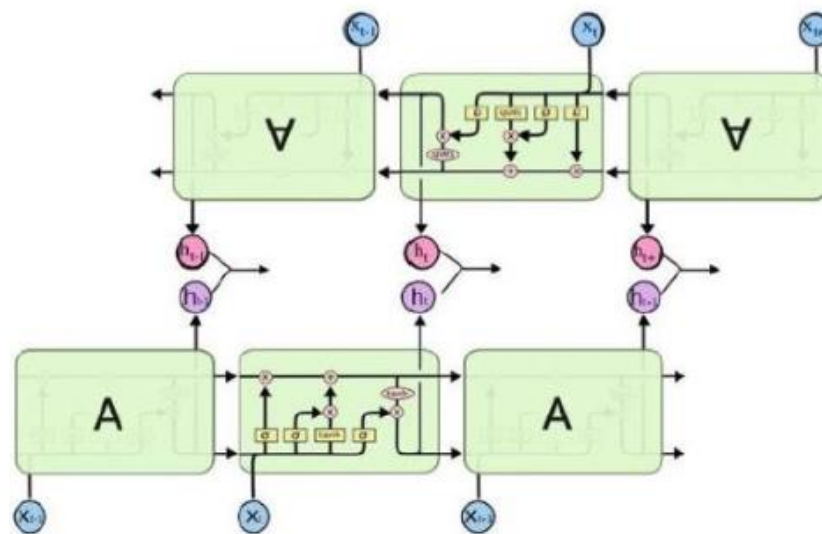
2.8 Bidirectional Long Short-Term Memory (BiLSTM)

Dengan adanya perkembangan untuk setiap model, *LSTM* juga mengalami perkembangan yaitu dengan menggabungkan dua lapisan *LSTM* yang sekarang disebut dengan *Bidirectional Long-Short Term Memory* (BiLSTM). Kombinasi dari dua lapisan *LSTM* dengan memproses rangkaian dari awal ke akhir (*forward LSTM*) dan memproses rangkaian dari akhir ke awal (*backward LSTM*), dengan dua lapisan itu BiLSTM dapat mengambil informasi konteks dari kedua penghujung rangkaiannya. Dengan begitu, proses pembelajaran menjadi lebih mendalam, yang berdampak pada kualitas model (Graves, et.all, 2013).



Gambar 2.16 *Arsitektur BiLSTM*
(Lu Peng, et all, 2020)

Pada Gambar 2.16 Menunjukkan bahwa representasi input data yang diberikan ke model adalah x_{n-1} , x_n , x_{n+1} . Masuk ke dalam lapisan *forward LSTM* dan data inputan akan diproses dari awal ke akhir. Setiap tahap waktu akan menghasilkan status tersembunyi dan outputnya, seperti waktu n , lapisan *forward LSTM* akan menghasilkan h_n dan S_n^f . Seperti pada tahapan *forward* data inputan juga masuk ke dalam lapisan *backward LSTM* dan memproses data dari akhir ke awal. Yang setiap tahapan waktunya menghasilkan h_n dan S_n^b . Output dari tahapan waktu yang dihasilkan pada lapisan *forward* dan *backward LSTM* akan digabungkan melalui fungsi aktivasi (σ atau \tanh) untuk menghasilkan output akhir. y_{n-1} , y_n , y_{n+1} merupakan output yang dihasilkan dari proses model *BiLSTM* pada setiap tahapan waktunya.



Gambar 2.17 *Model BiLSTM*

Sumber: Varsamopoulos, 2018

Gambar 2.17 menunjukkan model dari *BiLSTM* yang menggabungkan dua *LSTM forward layer* dan *backward layer*. *Hidden layer* juga terhung ke dalam *LSTM forward* dan *LSTM backward*.

Output dari LSTM dua arah *hidden layer* akan dijabarkan pada persamaan 2.9.

$$y_t = W_{hy}^{\rightarrow} \vec{h}_t + W_{hy}^{\leftarrow} \overleftarrow{h}_t \quad (2.9)$$

Keterangan:

y_t = *output gate LSTM* dua arah

W_{hy}^{\rightarrow} = nilai bobot untuk *output gate LSTM forward*

\vec{h}_t = nilai keluaran *LSTM forward*

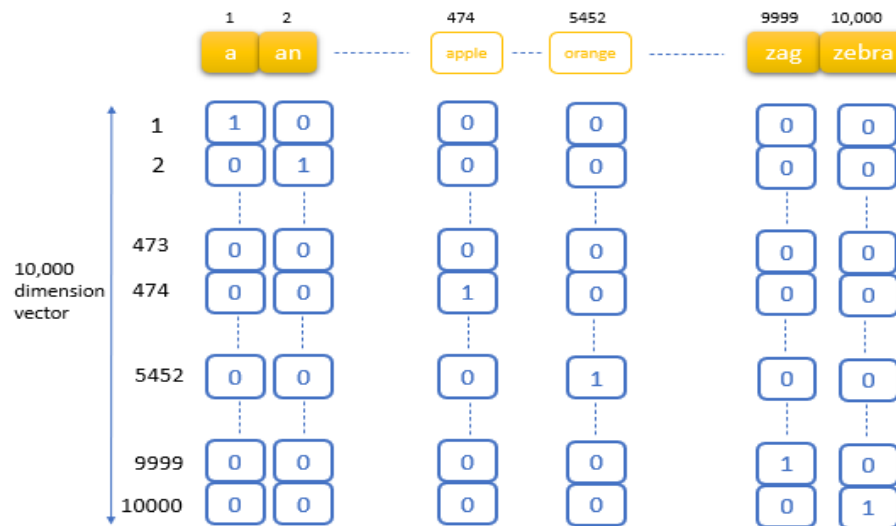
W_{hy}^{\leftarrow} = nilai bobot untuk *output gate LSTM backward*

\overleftarrow{h}_t = nilai keluaran *LSTM backward*

2.9 Word Embedding

Metode pembelajaran yang dikenal *word embedding* melibatkan penggabungan kata-kata dari kosa kata ke dalam bentuk *vektor numerik* (bilangan real). *Word embedding* diperlukan karena komputer dan algoritma AI tidak dapat memahami dan memproses kata seperti manusia. Kata-kata tersebut harus direpresentasikan dalam bentuk angka atau vektor agar dapat diproses oleh komputer dan algoritma AI (Mulyawan, 2023).

Awalnya, *one-hot encoding* bisa dianggap menjadi salah satu cara paling sederhana dalam merepresentasikan kata secara numerik. *One-hot encoding* berarti vektor yang terdiri dari satu elemen dengan nilai 1, dan elemen sisanya memiliki nilai 0, sehingga sebagian besar elemen dalam matriks kecuali diagonal utama memiliki nilai 0 dalam matriks (ini adalah indeks unik setiap kata) yang nilainya 1. Pada Gambar 2.18 merupakan contoh penerapan *One-hot encoding* untuk jumlah 10.000 kata:



Gambar 2.18 One-hot Encoding kata dalam kosa kata

Sumber: [raviteja-ganta.github.io](https://github.com/raviteja-ganta) 2021

Menggunakan *one-hot encoding* akan tidak efisien untuk melakukan konversi, karena bertambah panjang kalimat atau bertambah banyak kata, maka akan bertambah besar pula ukuran matriksnya. Bahkan untuk ukuran yang relatif kecil, teks akan memerlukan ruang penyimpanan yang besar secara eksponensial dan dengan demikian menjadi tidak efisien secara komputasi. Masalah lain dengan menggunakan *one-hot encoding* adalah bahwa semua kata memiliki spasi yang sama, yang membuatnya sangat sulit untuk menangkap hubungan antar kata, sehingga perlu menggunakan model untuk *embedding*. Terdapat beberapa model yang dapat digunakan seperti *Word2Vec*, *GloVe*, *BERT*, *FastText*, dll.

Dipenelitian ini metode yang diterapkan untuk tahapan pengkonversian kata menggunakan model *FastText*, yang merupakan peningkatan dari *library Word2Vec* yang telah dikenal sebagai metode untuk *word embedding*. Berbeda dengan *Word2Vec*, *FastText* menggunakan n-gram untuk memproses kata dan bukan hanya satu kata secara utuh saja. Contoh implementasi n-gram untuk kata “gendut” dengan trigram ($n = 3$) adalah “gen”, “end”, “ndu”, “dut”. Selain itu kelebihan yang dimiliki *FastText* ini adalah kecepatan dalam waktu memproses kata.

2.10 Penelitian Terdahulu

Penelitian terkait komentar *hate speech* telah dilakukan oleh Pasaribu (2022) menggunakan algoritma *Backpropagation Artificial Neural Network*. Penelitian ini menggunakan 1.550 data dan menghasilkan akurasi sebesar 87,1%, presisi sebesar 89,66%, dan recall sebesar 83,87%.

Penelitian lain terkait komentar *body shaming* di media sosial dilakukan oleh Pardede (2022) dengan algoritma *Term Frequency Chi-Square* dan *Random Forest Classifier*. Menggunakan 4.300 data yang diambil dari akun YouTuber Indonesia, penelitian ini menerapkan metode evaluasi cross-validation dan memperoleh akurasi sebesar 91,60%.

Berbagai penelitian telah dilakukan untuk mengklasifikasikan *cyberbullying*. Fadli dan Hidayatullah (2021) menggunakan algoritma *Long Short-Term Memory* (LSTM) dan *Bidirectional Long Short-Term Memory* (BiLSTM) untuk mengidentifikasi *cyberbullying* di Twitter berbahasa Indonesia. Menggunakan 6.835 cuitan yang dibagi menjadi cuitan *non-cyberbullying* dan *cyberbullying*. Hasil penelitian menunjukkan akurasi sebesar 93,77% untuk LSTM dan 95,24% untuk BiLSTM.

Penelitian lain oleh Raudhoti, Herdiani, dan Romadhony (2020) menggunakan algoritma *Support Vector Machine* (SVM) dan *Semantic Similarity* untuk mengidentifikasi *cyberbullying* pada kolom komentar Instagram. Dengan menggunakan 685 komentar, hasil pengujian menunjukkan bahwa penggunaan informasi *Semantic Similarity* meningkatkan akurasi klasifikasi sebesar 7%, dari 67% menjadi 74%.

Candra dan Rozana (2020) mengklasifikasikan komentar *bullying* di Instagram menggunakan algoritma *K-Nearest Neighbor* (KNN). Dari 1.000 data yang terdiri dari 500 komentar *bullying* dan 500 *non-bullying*, penelitian ini menghasilkan presisi rata-rata sebesar 58,83% dengan akurasi tertinggi mencapai 77%.

Syarif, Herdiani, dan Astuti (2019) mengidentifikasi *cyberbullying* pada komentar Instagram dalam konteks Pemilihan Presiden Indonesia tahun 2019. Dari 1.434 data yang terdiri dari 179 komentar *cyberbullying* dan 181 komentar *non-cyberbullying*, hasil penelitian menunjukkan bahwa performa *Lexicon-Based* mendapat akurasi sebesar 58%, sementara *Naïve Bayes Classifier* mencapai akurasi 97%.

Penelitian oleh Purnamasari, Fauzi, Indriati, dan Dewi (2018) mengidentifikasi tweet *cyberbullying* di Twitter menggunakan algoritma *Support Vector Machine* (SVM) dan *Information Gain* (IG). Dari 300 tweet yang terdiri dari 150 tweet *bully* dan 150 tweet *non-bully*, SVM mencapai akurasi sebesar 75%, sementara IG mencapai akurasi 76,66%.

Luqyana, Cholissodin, dan Perdana (2018) menggunakan algoritma *Klasifikasi Support Vector Machine* (SVM) untuk menganalisis sentimen *cyberbullying* pada komentar Instagram. Dari 400 data yang dikumpulkan, hasil pengujian menunjukkan akurasi tertinggi sebesar 90% pada komposisi data *testing* 50% dan data *training* 50%.

Ismiati (2018) menggunakan algoritma *Naïve Bayes Classifier* untuk mendeteksi komentar negatif di Instagram. Dari sekitar 50 data komentar Instagram yang dikumpulkan, penelitian ini mencapai akurasi sebesar 76,6%, dengan 23 dari 30 pengguna menghasilkan kategori komentar yang sesuai dengan yang diharapkan.

Tabel 2.4 Penelitian Terdahulu

No	Penulis	Judul	Metode	Keterangan
1.	Pasaribu. (2022)	Identifikasi Hate <i>Speech</i> pada Media Sosial <i>Twitter</i> Menggunakan <i>Backpropagation Artificial Neural Network</i>	<i>Backpropagation Artificial Neural Network</i>	Hasil yang diperoleh dalam penelitian ini memperoleh nilai akurasi sebesar 87.1%, presisi sebesar 89.66%, dan <i>recall</i> sebesar 83,87%.
2.	Pardede. (2022)	Identifikasi Komentar <i>Body Shaming</i> pada Media Sosial Menggunakan <i>Term Frequency Chi-Square</i> dan <i>Random Forest Classifier</i>	<i>Term Frequency Chi-Square</i> dan <i>Random Forest Classifier</i>	Dalam penelitian ini, peneliti menggunakan metode evaluasi <i>cross validation</i> digunakan dengan memperoleh hasil akurasi sebesar 91.60%.

Tabel 2.4 Penelitian Terdahulu (Lanjutan)

No	Penulis	Judul	Metode	Keterangan
3.	Fadli, & Hidayatullah. (2021)	Identifikasi cyberbullying pada media sosial twitter menggunakan metode lstm dan bilstm	<i>Long Short-Term Memory</i> (LSTM) dan <i>Bidirectional Long Short-Term Memory</i> (BiLSTM)	Berdasarkan hasil penelitian menunjukkan keakuratan setiap algoritma ditulis dalam <i>Long Short-Term Memory</i> 93,77% dan <i>Bidirectional Long Short-Term Memory</i> 95.24%.
4.	Raudhoti, Herdiani, & Romadhony. (2020)	Identifikasi Cyberbullying pada Kolom Komentar Instagram dengan Metode Support Vector Machine dan Semantic Similarity	<i>Support Vector Machine</i> (SVM) dan <i>Semantic Similarity</i>	Hasil dari pengujian menunjukkan bahwa penggunaan informasi <i>Semantic Similarity</i> meningkatkan akurasi klasifikasi sebesar 7%, dari 67% menjadi 74%.
5.	Candra, & Rozana. (2020)	Klasifikasi Komentar Bullying pada Instagram Menggunakan Metode K-Nearest Neighbor	<i>K-Nearest Neighbor</i>	Hasil dari pengujian menghasilkan akurasi dengan rata-rata sebesar 58,83%. Dan terdapat akurasi tertinggi yaitu 77%.

Tabel 2.4 Penelitian Terdahulu (Lanjutan)

No	Penulis	Judul	Metode	Keterangan
6.	Syarif, Herdiani, & Astuti. (2019)	Identifikasi Cyberbullying Pada Komentar Instagram Menggunakan Metode Lexicon-based Dan Naïve Bayes Classifier (studi Kasus: Pemilihan Presiden Indonesia Tahun 2019)	<i>Lexicon-Based</i> dan Naïve Bayes Classifier	Hasil dari penelitian ini mendapat performasi dari <i>Lexicon-Based</i> mendapat akurasi sebesar 58%. Sementara itu, Naïve Bayes Classifier mendapatkan akurasi sebesar 97%.
7.	Purnamasari, et all. (2018)	Identifikasi Tweet Cyberbullying pada Aplikasi Twitter menggunakan Metode Support Vector Machine (SVM) dan Information Gain (IG) sebagai Seleksi Fitur	<i>Support Vector Machine</i> (SVM) dan <i>Information Gain</i> (IG)	Hasil dari penelitian ini didapatkan <i>Support Vector Machine</i> (SVM) memiliki akurasi sebesar 75%. Sementara itu, <i>Information Gain</i> (IG) mendapatkan akurasi sebesar 76,66%.

Tabel 2.4 Penelitian Terdahulu (Lanjutan)

No	Penulis	Judul	Metode	Keterangan
8.	Luqyana, Cholissodin, & Perdana. (2018)	Analisis Sentimen Cyberbullying Pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine	Klasifikasi <i>Support Vector Machine</i> (SVM)	Hasil dari pengujian ini mendapatkan akurasi tertinggi sebesar 90% pada komposisi data <i>testing</i> 50% dan komposisi data <i>training</i> 50%.
9.	Ismiati. (2018)	Deteksi Komentar Negatif Di Instagram Menggunakan Algoritma <i>Naïve Bayes Classifier</i>	<i>Naïve Bayes Classifier</i>	Hasil dari penelitian ini mendapatkan hasil akurasi sebanyak 76,6% dimana ada 23 pengguna dari 30 yang hasil kategorin komentarnya sesuai dengan yang diharapkan.

Terdapat perbedaan dalam penelitian ini dengan penelitian terdahulu seperti dijelaskan sebelumnya. Penelitian terdahulu terbatas pada model untuk mengklasifikasikan cuitan sebagai *cyberbullying* atau *non-cyberbullying* di jejaring sosial. Dalam penelitian ini, klasifikasi tersebut akan diperluas menjadi beberapa kategori cyberbullying, yaitu *flaming*, *sexism*, *body shaming* dan netral.

BAB 3

ANALISIS DAN PERANCANGAN SISTEM

3.1 Dataset

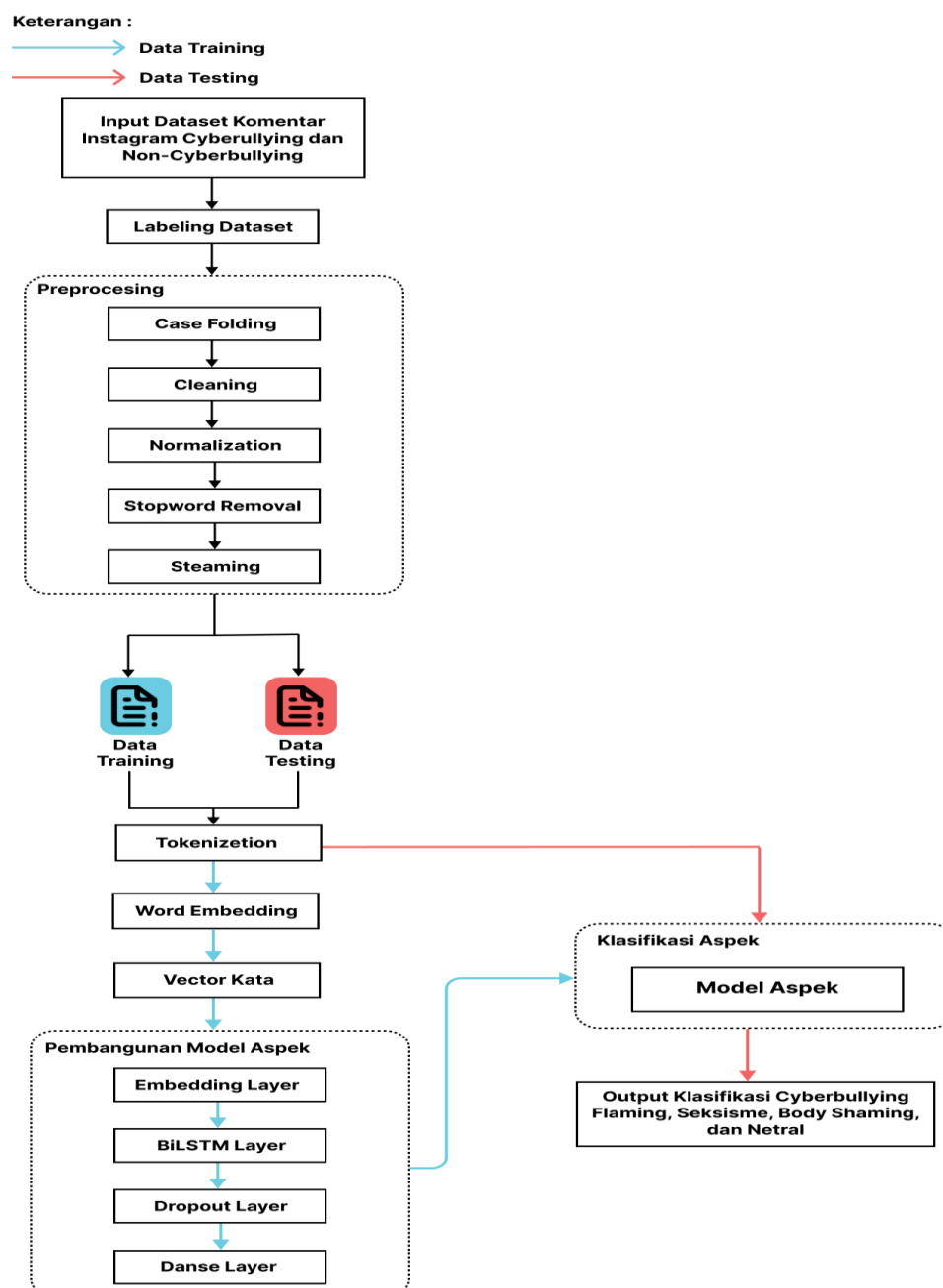
Dalam penelitian ini, penulis mengumpulkan dataset yang berisikan komentar berbahasa Indonesia, dengan inputan yang berupa teks yang dikumpulkan melalui aplikasi Instagram dengan menggunakan *tools* dari *library instaloader*. Hasil dari *crawling* data terdapat 8000 komentar yang mana akan disimpan ke dalam file *csv*. Table 3.1 menunjukkan hasil *crawling* data dari aplikasi Instagram.

Tabel 3.1 Hasil Crawling Komentar Instagram

NO.	KOMENTAR
1.	Karma baik akan mendatangkan kebaikan. Sebaliknya karma buruk akan membawa kau kedalam kehancuran kelak!! INGAT ITU YAH BINATANG
2.	Perempuan gak bisa menangani pekerjaan yang membutuhkan ketegasan dalam pengambilan keputusan.
3.	Alay norak lu Robby anjng, sok sok penyelamat eh malah kena hujat se Indonesia
4.	Darah nya jadi darah kotor daging nya jadi daging haram itulah isi perut bapak ini
5.	dgn diviralin gini tanpa konfirmasi atau samperin aja langsung bapaknya bilangin langsung,, dibalik dia kerja dia harus hidupin keluarganya kan mba,, yg gajinya jg pasti jauh dr pendapatan mbaknya,, dgn viral begini dia udh dipecat apakah msh ada vendor yg mau terima dia kerja mbak,, iya mbak orang berada n dog person makanya tanpa samperin langsung bapaknya,,mbak langsung bela bgt dognya ya mba
6.	Denger suaranya aja langsung pusing..liat muka nya langsung pengen muntah
7.	Dihhhh pahanya gede banget, klo udh gitu banyakin olahraga sih.

3.2 Arsitektur Umum

Dalam penelitian ini, Gambar 3.1 menunjukkan struktur rancangan umum, yang kerja sistemnya dibagi menjadi beberapa tahapan, dari pengumpulan data dari aplikasi Instagram. Data yang telah dikumpulkan akan masuk ke tahap *preprocessing*, yang meliputi tahapan *case folding*, *cleaning*, *normalization*, *stopword removal*, *steaming*, dan *tokenization*. Selanjutnya data dibagi menjadi data latih dan data uji, dan setiap kata diubah membentuk matriks melalui tahapan *word embedding* yang siap untuk digunakan dalam tahap pelatihan. Setelah itu, model akan dibuat untuk mengklasifikasikan aspek komentar Instagram, yang kemudian akan dikategorikan ke dalam komentar *cyberbullying* dan *non-cyberbullying*.



Gambar 3.1 Arsitektur Umum

3.2.1 Input Dataset

Pada tahapan ini, dataset komentar *cyberbullying* dan *non-cyberbullying* yang dipilih secara acak dari beberapa pengguna Instagram dikumpulkan. Data tersebut diperoleh dengan mengumpulkan data dari aplikasi Instagram yang menggunakan *tools* dari *library instaloader* untuk mengumpulkan *tag*, foto, *feeds* dan komentar di Instagram. Dataset ini terdiri dari empat *class* yang akan dikelompokkan yaitu *class 1 sexism*, *class 2 flaming*, *class 3 body-shaming*, dan *class 4 netral*.

Proses pelebelan data dilakukan secara manual oleh dua ahli, Jihan Septiani Hasan S.Pd, seorang guru bahasa, dan Deby Saphira S.Pd, seorang ahli konseling, untuk menilai apakah komentar tersebut termasuk dalam kategori *class* tersebut atau tidak berdasarkan karakteristik dan kata kunci yang telah ditetapkan pada Bab 2. Setelah itu data yang telah diperoleh akan disimpan data dalam sebuah file dengan *identifier csv* (*Comma Seperated Value*). Table 3.2 berisikan data komentar dari aplikasi Instagram yang sudah dilabeli.

Tabel 3.2 Komentar Yang Dilebeli

<i>No.</i>	Comment	<i>Label</i>
1	Pinggangmu terlalu lebar, coba kurangi asupan kalori.	3
2	Kenapa kamu tidak pakai baju yang menutupi lengan?	3
3	Dasar orang nggak punya otak, ngomongnya aja udah bikin kesel.	2
4	Kenapa kamu tidak malu dengan bentuk tubuhmu?	3
5	Laki-laki tidak perlu belajar seni patung.	1
...
7996	Laki-laki lebih berhak untuk menentukan arah dalam hubungan	1
7997	Wanita tidak bisa memahami ilmu komputer sebaik laki-laki.	1
7998	Bacott lo tu nggak paham apa-apa tentang hidup.	2
7999	Lokasinya cantik sekali...	4
8000	Pemandangan yang menakjubkan.	4

Terdapat pembagian pada dataset, dimana data latih dan data uji dengan parameter 80:20 dari hasil *crawling* data pada aplikasi Instagram. Terdapat 6400 data *training* dan 1600 data *testing*.

3.2.2 *Pre-processing*

Preprocessing adalah tahap penting yang melibatkan akumulasi data mentah yang kemudian diolah menjadi informasi yang berguna. Data mentah harus diproses dikarenakan masih mengandung data yang *incomplete* (kurang atribut), *noise* (masih mengandung *error* dan *outliers*) dan *inconsistent* (bertentangan dengan penamaan, singkatan, dan kode). Oleh karena itu, *preprocessing* ini menjadi langkah krusial dalam analisis klasifikasi, yang berfungsi untuk membersihkan data dari faktor-faktor yang tidak diperlukan untuk mempersingkat proses pengklasifikasian. Berikut ini merupakan tingkatan dari *pre-processing* pada penelitian ini:

a. *Case Folding*

Tahapan ini berguna untuk menggantikan semua huruf besar dalam teks dengan huruf kecil (normal). Proses ini hanya mencakup huruf “a” sampai “z” yang diproses. Hal ini dilakukan untuk mempermudah pengolahan dan meningkatkan kepekaan terhadap penulisan kalimat. Tahapan *cace folding* pada Table 3.3.

Tabel 3.3 Penerapan *Cese Folding*

Sebelum <i>Case Folding</i>	Sesudah <i>Case Folding</i>
Tanggung jawab lo woiii..! Sok Viralin Bkin org kehilangan kerjaan.. mknya Klo Keluar Rumah OTAK LO D BAWA MBAAKKKK!!!	tanggung jawab lo woiii..! sok viralin bkin org kehilangan kerjaan.. mknya klo keluar rumah otak lo d bawa mbaakkkk!!!

b. *Cleaning*

Tahapan ini menghilangkan karakter khusus, membersihkan kata dan menghilangkan tanda bacaan seperti koma(,), titik(.), menghilangkan angka, menghilangkan karakter-karakter atau simbol-simbol, menghilangkan URL (*Uniform Resource Locator*) serta data kosong. Karena proses pendataan komentar Instagram secara mekanis tidak selalu berjalan sempurna dan juga mengurangi *noise* maka proses ini penting dilakukan. Tahapan *cleaning* pada Table 3.4.

Tabel 3.4 Penerapan Cleaning

Sebelum <i>Cleaning</i>	Sesudah <i>Cleaning</i>
tanggung jawab lo woiii..! sok viralin bkin org kehilangan kerjaan.. mknya klo keluar rumah otak lo d bawa mbaakkkk!!!	tanggung jawab lo woiii sok viralin bkin org kehilangan kerjaan mknya klo keluar rumah otak lo d bawa mbaakkkk

c. *Normalization*

Komentar di media sosial Instagram biasanya ditulis dalam bentuk teks yang tidak mengikuti standart penulisan yang benar, mulai dari singkatan kata, istilah modern, bahasa daerah, dan istilah asing. langkah normalisasi ini mengubah format teks dari non-formal menjadi formal untuk mempermudah proses dengan menggunakan kamus bahasa gaul yang tersedia. Sebagai contoh kata “gk” akan dinormalisasikan menjadi “tidak”, kata “aaahhh” dinormalisasikan menjadi “ah” dan lain sebagainya. Namun kamus yang digunakan juga mempunyai kekurangan, artinya terkadang masih ada kata yang tidak dapat dinormalisasikan karena kesalahan penulisan dan singkatan yang berlebihan atau baru yang ada dalam kamus. Table 3.5 berikut menunjukkan tahapan dari *normalization*.

Tabel 3.5 Penerapan Normalization

Sebelum <i>Normalization</i>	Sesudah <i>Normalization</i>
tanggung jawab lo woiii sok viralin bkin org kehilangan kerjaan mknya klo keluar rumah otak lo d bawa mbaakkkk	tanggung jawab kamu woi berlagak viralin buat orang kehilangan kerjaan makanya kalau keluar rumah otak kamu di bawa mbak

d. *Stopword Removal*

Setelah proses lainnya, teks akan diproses pada tahapan *stopword removal*. *Stopword* adalah kata-kata yang dianggap tidak penting dalam teks komentar, seperti preposisi, pronomina, konjungsi dan lain-lain. Langkah ini memiliki tujuan dalam mengurangi dimensi data yang akan menghapus setiap kata yang tidak memiliki nilai informatif dengan begitu dapat membantu meningkatkan efisiensi dan akurasi dari model tersebut. Kata-kata tersebut mencakup “dan”, “di”, “atau”, “pada”, “ke”, “dari”, “yang” dan sebagainya. Metode yang digunakan dalam proses *stopword removal* ini adalah berdasarkan pustaka (*library*) *Sastrawi*. Table 3.6 berikut menunjukkan tahapan dari *stopword removal*.

Tabel 3.6 Penerapan Stopword Removal

Sebelum <i>Stopword Removal</i>	Sesudah <i>Stopword Removal</i>
tanggung jawab kamu woi berlagak viralin buat orang kehilangan kerjaan makanya kalau keluar rumah otak kamu di bawa mbak	tanggung jawab berlagak viralin kehilangan kerjaan keluar rumah otak di bawa

e. *Steaming*

Dalam proses pengolahan teks, *steaming* bertujuan untuk mendapatkan kata dasar dari kata yang memiliki imbuhan. Pada tahapan ini, semua imbuhan termasuk *prefix* (awalan), *infix* (sisipan), *suffix* (akhiran), dan *conffix* (gabungan) akan dihilangkan. Adapun contoh dari kata yang menggunakan imbuhan seperti “membaca” menjadi kata dasar “baca”, kata “jelasnya” menjadi “jelas”, kata “temali” menjadi “tali”, dan kata “berjatuhan” menjadi “jatuh” dan masih banyak lagi contoh lainnya. Table 3.7 berikut menunjukkan tahapan dari *stopword steaming*.

Tabel 3.7 Penerapan Steaming

Sebelum <i>Steaming</i>	Sesudah <i>Steaming</i>
tanggung jawab berlagak viralin kehilangan kerjaan keluar rumah otak di bawa	tanggung jawab lagak viral hilang kerja keluar rumah otak bawa

f. *Tokenization*

Tokenisasi adalah proses pemecahan kalimat panjang menjadi bagian-bagian yang lebih kecil, seperti paragraph, kalimat dan kata (*token*). Tahap tokenisasi juga dikenal segmentasi teks atau analisis leksikal. Dalam konteks ini, tokenisasi merupakan pemecahan kalimat menjadi kata-kata penyusunnya. Proses ini terlihat mudah, namun sebenarnya proses ini dangat sulit. Tanda baca atau spasi dapat digunakan untuk memisahkan kata dan kalimat. Artinya, jika dua kata digabungkan dengan spasi, kata-kata tersebut dipisahkan oleh spasi. Table 3.8 berikut menunjukkan tahapan dari *tokenization*.

Tabel 3.8 Penerapan Tokenization

Sebelum <i>Tokenization</i>	Sesudah <i>Tokenization</i>
tanggung jawab lagak viral hilang kerja keluar rumah otak bawa	“tanggung”, “jawab”, “lagak”, “viral”, “hilang”, “kerja”, “keluar”, “rumah”, “otak”, “bawa”

3.2.3 *Word Embedding*

Pada tahap *word embedding*, setiap kata akan diwakili dengan *vector* yang bernilai nyata dalam ruang *vector* tertentu. Tahapan ini sering dikaitkan dengan pembelajaran mendalam karena masing-masing dari kata akan dimuat ke satu *vector*, dan nilai *vector* tersebut akan dianalisis dengan cara yang sama dengan jaringan saraf dalam kecerdasan buatan (AI). *Fasttext* adalah salah satu *library* yang diterapkan dalam penelitian ini dapat menerapkan konsep dari *word embedding*. Pada langkah ini, setiap kata yang dihasilkan dari tokenisasi akan diubah menjadi matriks *vector*, yang akan berfungsi sebagai bobot (*weight*) pada jaringan saraf tiruan. Gambar 3.2 menunjukkan contoh dari *word embedding matrix*, di mana *vector* dari kata “bodoh”.

```

array([ 0.0121 ,  0.0058 , -0.0452 ,  0.1015 , -0.075 ,
       -0.1247 , -0.0133 ,  0.0249 ,  0.0154 , -0.1051 ,
       -0.0728 , -0.0359 ,  0.0269 ,  0.0065 ,  0.0234 ,
        0.0126 ,  0.0215 ,  0.0126 , -0.0186 ,  0.0599 ,
       -0.0307 , -0.0431 , -0.0052 ,  0.0114 ,  0.0292 ,
       -0.0421 , -0.1239 , -0.0313 ,  0.0338 , -0.0214 ,
        0.0448 , -0.0358 , -0.0189 , -0.1318 ,  0.0346 ,
       -0.0309 ,  0.0593 ,  0.115 , -0.0106 ,  0.0598 ,
       -0.0099 ,  0.0513 , -0.0004 ,  0.0333 , -0.028 ,
        0.0535 , -0.0503 ,  0.0939 , -0.0052 , -0.0158 ,
        0.0394 ,  0.0547 , -0.0741 , -0.0035 ,  0.0058 ,
        0.0348 , -0.0536 , -0.0106 , -0.1021 ,  0.0174 ,
       -0.003 , -0.0193 , -0.0437 , -0.0252 ,  0.0475 ,
        0.0012 , -0.0244 ,  0.0978 ,  0.1107 ,  0.0731 ,
        0.0248 ,  0.036 ,  0.0287 ,  0.0059 , -0.0657 ,
        0.0498 ,  0.0278 ,  0.0303 , -0.0397 ,  0.0541 ,
        0.07 ,  0.0278 ,  0.0244 , -0.0686 ,  0.025 ,
       -0.0068 , -0.0978 , -0.0536 , -0.0162 ,  0.0239 ,
        0.0067 , -0.0288 ,  0.0011 ,  0.0107 , -0.17290001,
       -0.1293 ,  0.0378 , -0.0298 , -0.0022 ,  0.0575 ,
        0.0008 , -0.0324 , -0.015 , -0.0213 ,  0.046 ,
        0.0524 , -0.1425 ,  0.053 ,  0.02 , -0.0522 ,
       -0.022 , -0.1013 ,  0.0723 , -0.0654 , -0.0277 ,
       -0.0308 , -0.07 , -0.0809 ,  0.0638 ,  0.0219 ,
       -0.067 , -0.0363 ,  0.054 ,  0.06 ,  0.0532 ,
        0.0511 ,  0.0362 , -0.0372 , -0.0139 , -0.0008 ,
        0.0226 , -0.0558 ,  0.0968 ,  0.0243 ,  0.0297 ,
       -0.0965 ,  0.03 ,  0.0109 ,  0.0304 ,  0.0882 ,
        0.1193 , -0.17110001, -0.006 ,  0.0079 ,  0.0405 ,
        0.0037 ,  0.0452 ,  0.0371 ,  0.0328 , -0.0183 ,
       -0.0281 , -0.0337 , -0.0721 ,  0.0086 ,  0.0607 ,
        0.024 ,  0.0081 ,  0.0516 ,  0.0061 , -0.0946 ,
        0.0913 ,  0.0696 , -0.0626 ,  0.0543 ,  0.0134 ,
       -0.064 ,  0.1004 ,  0.0076 ,  0.0027 ,  0.0069 ,
        0.0535 , -0.0384 ,  0.0336 , -0.0135 ,  0.049 ,
       -0.0436 ,  0.0004 ,  0.0605 ,  0.124 ,  0.002 ,
        0.1715 ,  0.0076 ,  0.0269 , -0.1205 ,  0.0557 ,
       -0.0103 ,  0.0795 , -0.014 ,  0.0227 , -0.1105 ,
       -0.0465 ,  0.0461 , -0.0208 ,  0.0081 , -0.0692 ,
       -0.0489 , -0.0266 ,  0.0362 ,  0.0478 ,  0.0258 ,
        0.0269 ,  0. ,  0.0155 , -0.0201 ,  0.0541 ,
        0.0209 ,  0.0016 ,  0.0415 , -0.002 ,  0.0568 ,
        0.14830001, -0.0338 ,  0.039 ,  0.0398 , -0.0723 ,
        0.0056 ,  0.0584 , -0.0304 , -0.0162 , -0.0737 ,
       -0.0515 ,  0.0578 , -0.0484 , -0.0066 , -0.0925 ,
        0.0334 , -0.0607 ,  0.0563 , -0.0287 , -0.0116 ,
       -0.0395 ,  0.0399 ,  0.014 , -0.0658 , -0.0136 ,
       -0.0312 , -0.0761 , -0.0379 ,  0.0784 ,  0.0323 ,
        0.038 ,  0.0768 ,  0.0256 , -0.0513 ,  0.001 ,
       -0.0033 , -0.012 , -0.0336 ,  0.0542 , -0.0197 ,
        0.0433 , -0.0244 ,  0.0264 ,  0.0482 ,  0.0154 ,
        0.055 , -0.0126 ,  0.0375 ,  0.0102 , -0.0457 ,
       -0.0235 ,  0.0279 , -0.0788 , -0.013 , -0.0927 ,
       -0.0578 , -0.016 ,  0.0041 , -0.0264 ,  0.0636 ,
        0.0769 ,  0.0146 , -0.0141 , -0.0537 ,  0.09 ,
        0.0164 , -0.062 , -0.0723 , -0.0733 , -0.0934 ,
       -0.0993 , -0.013 , -0.0983 , -0.0273 ,  0.0716 ,
        0.15880001, -0.0159 , -0.0484 , -0.0268 ,  0.0112 ,
       -0.0738 , -0.1476 , -0.0377 ,  0.0199 , -0.0818 ,
        0.0346 , -0.0255 , -0.0121 ,  0.0569 ,  0.0697 ]])

```

Gambar 3.2 Representasi Vector Berdasarkan Kata “bodoh”

3.2.4 Pembangunan Model

Proses selanjutnya yang akan dijalankan adalah pembangunan model dengan menerapkan algoritma *Bidirectional Long-Short Term Memory*. Dikarenakan algoritma ini terbukti sangat akurat dalam klasifikasi, data teks akan dimasukkan ke dalam kelas *cyberbullying* (Agrawal & Awekar, 2018). Model ini akan dilengkapi dengan beberapa lapisan yang bertujuan untuk menerima *input* dan mengecilkan kemungkinan terjadinya *overfitting* dan *underfitting*. Beberapa layer akan digunakan dalam penelitian ini:

a. *Embedding Layer*

Lapisan *embedding* adalah lapisan jaringan yang tersembunyi yang menangani urutan kata berukuran tetap. Vektor yang bernilai nyata, juga dikenal sebagai penyisipan kata. *Embedding layer* dapat digunakan untuk mengubah inputan kata atau kalimat menjadi vektor numerik. Ini juga dikenal sebagai *word embedding*. Lapisan *embedding* yang menerima input dari hasil tokenisasi dan kemudian menggunakan berat atau bobot dari *word embedding*. Proses *padding* telah dilakukan pada *input* yang dimasukkan ke *embedding layer* sebelumnya sehingga panjang setiap baris nomor *token* memiliki ukuran yang sama. Berdasarkan dimensi yang dihasilkan, maka terbentuklah vektor berdimensi 300 sebagai *output*. Penggalan *code* dari *embedding layer* pada Gambar 3.3.

```
model.add(Embedding(input_dim=vocab_size, output_dim=embedding_dim,
weights=[embedding_matrix], input_length=max_len, trainable=False))
```

Gambar 3.3 Penggalan Code Embedding Layer

(Brownlee, 2017) Lapisan tersembunyi pertama pada suatu jaringan disebut *embedding layer*. Lapisan ini memerlukan 3 argumen:

- ***Input_dim*** : adalah ukuran kosakata atau jumlah kata unik yang terdapat pada data teks. Misalnya, jika memiliki 1.000 kata unik dalam kosakata maka '*input_dim*' akan bernilai 1.000.
- ***Output_dim*** : adalah ukuran vektor embedding yang dihasilkan pada setiap kata. Misalnya, vektor memiliki dimensi 50 maka '*output_dim*' akan memiliki nilai 50.
- ***Input_length*** : adalah panjang rangkaian *input* yang akan ditetapkan untuk setiap lapisan *input* model keras.

Dan untuk melakukan perintah agar tidak mengubah setiap angka yang ada selama melatih model akan dijalankan perintah *trainable=False*.

b. *Bidirectional Long Short-Term Memory (BiLSTM) Layer*

Dengan menggabungkan dari kedua arah, lapisan depan (*forward*) menunjukkan konteks sebelumnya, dan lapisan belakang (*backward*) menunjukkan konteks setelahnya. Dengan menggunakan lapisan tersembunyi, Bidirectional LSTM dapat memproses informasi dari dua arah. Hal ini dapat dicapai dengan membungkus lapisan LSTM tersembunyi dengan lapisan *Bidirectional*, dijabarkan Gambar 3.4:

```
model.add(Bidirectional(LSTM(units=64, return_sequences=True)))
model.add(Bidirectional(LSTM(units=32)))
```

Gambar 3.4 Pemenggalan Code Bidirectional LSTM Layer

Hal ini akan menghasilkan dua salinan lapisan tersembunyi, satu lapisan akan sesuai dengan urutan input sebagaimana adanya, dan yang lainnya akan terletak pada salinan terbalik dari urutan input. Nilai dari output *LSTM* ini akan dirangkai secara default. Dengan *LSTM unit* 32, maka alih-alih *TimeDistributed layer* menerima 16 *timesteps* dari 32 *output*, maka sekarang akan menerima 16 *timesteps* dari 64 *output* (32 unit + 32 unit).

c. *Dropout Layer*

Pada layer ini memungkinkan menghentikan bagian-bagian tertentu (*neuron*) dalam sebuah lapisan dengan peluang tertentu. Pelatihan akan lebih cepat dengan menggunakan *dropout layer* yang menyetel setengah dari aktivasi lapisan menjadi nol. Lapisan *dropout* ini digunakan untuk mencegah overfitting dan mempercepat proses pembelajaran. Cara kerja lapisan ini adalah mengurangi kompleksitas model yang dibangun dengan *dropout* sebagian *nodes* secara acak. Pada penelitian ini akan men-*dropout* sebanyak 50% dari *nodes* yang ada. Penggalan *code* dari *dropout layer* Gambar 3.5.

```
model.add(Dropout(0.5))
```

Gambar 3.5 Pemenggalan Code Dropout Layer

d. *Dense Layer*

Lapisan ini dikenal juga sebagai lapisan *fully-connected* yang mana setiap jaringan neuronnya terhubung dengan setiap *neuron* pada lapisan selanjutnya. Untuk *neuron* yang terdapat didalam *dense layer* akan melangsungkan operasi linier dalam inputannya yang diikuti fungsi aktivasi (*non-linieritas*) yang umunya diputuskan pada outputnya. Lapisan ini memiliki parameter penting seperti *units*, *activation*, dan *input_shape*. Dalam tahapan ini, penulis menggunakan fungsi *softmax activation*

dalam *output layer* model jaringan saraf yang memprediksi distribusi probabilitas multinomial. Untuk masalah klasifikasi *multi-class* di mana keanggotaan *class* diperlukan untuk lebih dari dua label *class*, maka *softmax* digunakan sebagai fungsi aktivasi. Penulis akan mengklasifikasi empat *class* dalam penelitian, penggunaan *softmax* di lapisan *dense layer* pada Gambar 3.6.

```
model.add(Dense(4, activation='softmax'))
```

Gambar 3.6 Pemenggalan Code Dense Layer

Data harus dipersiapkan saat memodelkan masalah klasifikasi *multi-class*. Variable target yang mengandung label *class* dikodekan terlebih dahulu, setiap label *class* diwakili oleh bilangan bulat mulai 0 hingga N-1, yang mana N merupakan jumlah label *class*. Dalam penelitian ini memiliki empat label *class* akan dikodekan sebagai bilangan bulat 0, 1, 2, dan 3. Kemudian, akan dikodekan menjadi vektor:

- Class 0: [1, 0, 0, 0]
- Class 1: [0, 1, 0, 0]
- Class 2: [0, 0, 1, 0]
- Class 3: [0, 0, 0, 1]

Ini menunjukkan distribusi probabilitas multinomial yang diharapkan untuk setiap *class* yang digunakan untuk mengoreksi model *supervised learning*.

3.2.5 Melatih Model

Model yang telah dipelajari (*learned model*) adalah rancangan yang dapat menemukan pola yang sebelumnya tidak terlihat atau membuat keputusan berdasarkan kumpulan data. *Learned model* merupakan file yang dilatih untuk mengenali pola tertentu. Setelah proses pelatihan dan pembangunan model telah selesai dipelajari oleh sistem pada tahap sebelumnya, data testing yang telah dilakukan pada tahap pembangunan model digunakan untuk menguji model yang telah dilatih. Gambar 3.7 penggalan code dari *Learned model*.

```
history = model.fit(X_train_padded, y_train, epochs=10,
                    batch_size=32, validation_data=(X_test_padded, y_test))
```

Gambar 3.7 Penggalan Code Learned Model

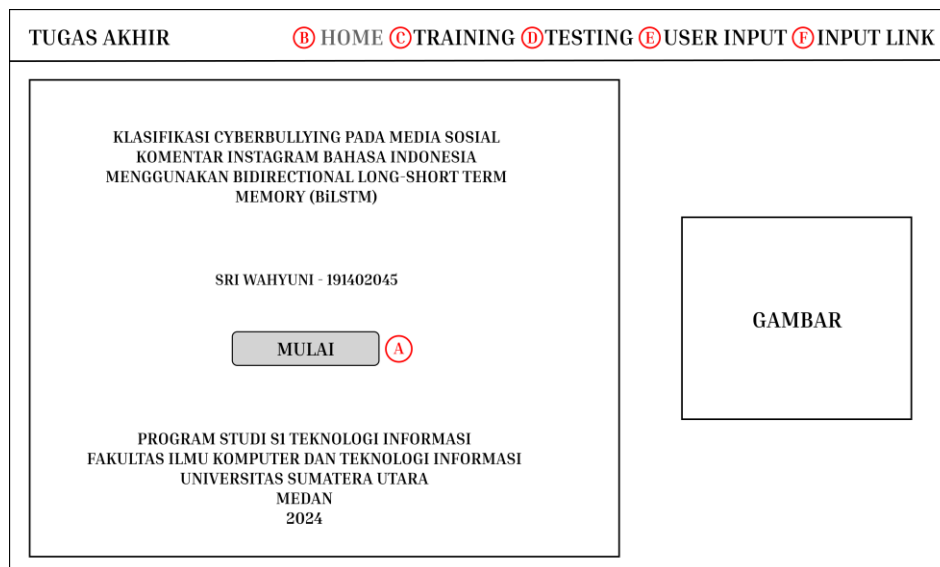
Dalam program memanggil fungsi *fit()*, yang digunakan untuk melatih model, dengan mengiris data menjadi “*batches*” dengan ukuran *batch_size* dan mengulangi seluruh kumpulan data untuk sejumlah yang diberikan *epochs*.

3.3 Perancangan Sistem

Bagian perancangan sistem berfungsi dalam merancang setiap alur proses dalam sistem ini, termasuk perancangan antarmuka sistem untuk aplikasi klasifikasi *cyberbullying* pada komentar di Instagram dengan menerapkan algoritma *Bidirectional Long-Short Term Memory* (BiLSTM). Sistem antarmuka akan dirancang berupa sistem web satu halaman (*one-page web*), yang bertujuan untuk memberikan gambaran umum seperti alur sistem dan tata letak menu yang mudah dipahami, sehingga membuat sistem mudah digunakan dan dioperasikan.

3.3.1 Rancangan Tampilan Bagian Beranda (*Home*)

Home page yang juga dikenal sebagai beranda merupakan halaman utama pada sistem yang akan ditampilkan ketika sistem dijalankan pertama kali. Halaman ini akan berisi judul dari penelitian, identitas dari penulis, serta navigasi ke bagian lainnya. Gambar 3.8 berikut menunjukkan desain tampilan *home page*.



Gambar 3.8 *Home Page (Rancangan)*

Keterangan:

1. Label (A) adalah *button* Mulai berfungsi untuk mengarahkan ke halaman *training*.
2. Label (B) menu *Home* berfungsi untuk menampilkan dan mengarahkan ke halaman awal sistem.
3. Label (C) menu *Training* berfungsi untuk mengarahkan ke halaman *training*.
4. Label (D) menu *Testing* berfungsi untuk mengarahkan ke halaman *testing*.
5. Label (E) menu *User Input* berfungsi untuk mengarahkan ke halaman mengklasifikasi jenis komentar secara *realtime*.
6. Label (F) menu *Input Link* berfungsi untuk mengarahkan ke halaman pengklasifikasian jenis komentar dengan menginput *link* postingan di Instagram.

3.3.2 Rancangan Tampilan Bagian *Training*

Dalam halaman *training* akan mencakup tampilan berupa *form* yang memungkinkan pengguna mengunggah file *.csv* yang disediakan dan menjalankan proses pelatihan model. Setelah file diunggah, pengguna dapat melanjutkan proses *training*. Jika proses *training* selesai maka akan menampilkan hasil dari *preprocessing* ditampilkan pada salah satu teks inputan, lalu akan menampilkan nilai *loos* dan *accuracy* dari data *training*. Rancangan dari *training page* ditunjukkan Gambar 3.9.

Gambar 3.9 Training Page (Rancangan)

Keterangan:

1. Label (A) adalah *button input* berfungsi untuk memasukkan file dataset berformat *.csv* untuk melakukan proses *training*.
2. Label (B) *button submit* berfungsi untuk menjalankan proses *training*.
3. Label (C) berfungsi untuk menampilkan hasil *preprocessing* dari salah satu data *training*.
4. Label (D) akan menampilkan grafik *loss* dari hasil *training*.
5. Label (E) akan menampilkan grafik *accuracy* dari hasil *training*.

3.3.3 Rancangan Tampilan Bagian *Testing*

Halaman *testing* akan menampilkan sebuah *form* untuk menerima berkas berekstensi *.csv* yang diunggah pengguna, yang akan digunakan untuk memulai proses pengujian (*testing*). Setelah proses pengujian selesai, hasil prediksi dari setiap aspek akan ditampilkan berbentuk tabel, seperti yang ditunjukkan dalam Gambar 3.10.

TUGAS AKHIR	HOME TRAINING TESTING USER INPUT INPUT LINK												
<div> TESTING Input File Testing Berekstensi .csv </div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;"> <div style="display: flex; align-items: center;"> <div style="border: 1px solid black; padding: 2px 5px; margin-right: 5px;"> A Pilih File </div> <div style="border: 1px solid black; flex-grow: 1; padding: 2px;">No file chosen</div> </div> <div style="display: flex; align-items: center; margin-top: 5px;"> <div style="background-color: black; color: white; padding: 5px 15px; border: 1px solid black; margin-right: 5px;"> Submit </div> <div style="border: 1px solid red; border-radius: 50%; padding: 0 2px; margin-left: 5px;">B</div> </div> </div> <div style="text-align: center; margin-top: 20px;"> C Confussion Matrix <div style="border: 1px solid black; width: 150px; height: 100px; margin: 10px auto; position: relative;"> <div style="position: absolute; top: 0; bottom: 0; left: 0; right: 0; border-left: 1px solid black; border-right: 1px solid black; border-top: 1px solid black; border-bottom: 1px solid black;"></div> <div style="position: absolute; top: 0; right: 0; bottom: 0; left: 0; border-left: 1px solid black; border-right: 1px solid black; border-top: 1px solid black; border-bottom: 1px solid black;"></div> </div> </div> <div style="margin-top: 10px;"> D HASIL KLASIFIKASI KOMENTAR BERDASARKAN ASPEK <div style="border: 1px solid black; padding: 5px; margin-top: 5px;"> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td style="height: 20px;"></td><td style="height: 20px;"></td><td style="height: 20px;"></td></tr> <tr><td style="height: 20px;"></td><td style="height: 20px;"></td><td style="height: 20px;"></td></tr> <tr><td style="height: 20px;"></td><td style="height: 20px;"></td><td style="height: 20px;"></td></tr> <tr><td style="height: 20px;"></td><td style="height: 20px;"></td><td style="height: 20px;"></td></tr> </table> </div> </div>													

Gambar 3.10 *Testing Page (Rancangan)*

Keterangan:

1. Label (A) *button input* berfungsi untuk memasukkan file dataset berformat .csv untuk melakukan proses *testing*.
2. Label (B) *button submit* berfungsi untuk menjalankan proses *training*.
3. Label (C) berfungsi untuk menampilkan hasil dari *confussion matrix*.
4. Label (D) berfungsi untuk menampilkan hasil prediksi berdasarkan aspek yang terdapat di data *training*.

3.3.4 Rancangan Tampilan Bagian *User Input*

Laman *user input* menggambarkan halaman di mana pengguna dapat menganalisis data tunggal. Pada halaman ini, sistem dapat mengklasifikasikan inputan yang terdiri dari teks yang diberikan oleh pengguna, dan menghasilkan prediksi komentar dan aspek berdasarkan komentar yang diberikan oleh pengguna. Gambar 3.11 menunjukkan rancangan tampilan *user input*.

TUGAS AKHIR	HOME	TRAINING	TESTING	USER INPUT	INPUT LINK
-------------	------	----------	---------	------------	------------

USER INPUT - UJI COBA REALTIME

Masukkan Komentar Menggunakan Bahasa Indonesia Untuk Diklasifikasi

A

B Submit

C **HASIL PREDIKSI REALTIME**

Gambar 3.11 *User Input Page (Rancangan)*

Keterangan:

1. Label (A) *input text* berfungsi untuk memasukkan teks komentar untuk melakukan prediksi aspek terhadap komentar yang diinput.
2. Label (B) *button submit* berfungsi untuk menjalankan proses klasifikasi.
3. Label (C) berfungsi untuk menampilkan hasil prediksi komentar berdasarkan aspek yang diinput pengguna.

3.3.5 Rancangan Tampilan Bagian Input Link

Halaman *input link* merupakan sistem yang dapat mengklasifikasikan inputan yang terdiri dari *link* yang diberikan oleh pengguna, dan menghasilkan prediksi komentar dan aspek berdasarkan komentar yang diberikan oleh pengguna. Gambar 3.12 menunjukkan rancangan tampilan *input link*.

TUGAS AKHIR	HOME	TRAINING	TESTING	USER INPUT	INPUT LINK
-------------	------	----------	---------	------------	------------

INPUT LINK POSTINGAN INSTAGRAM

Masukkan Link Postingan Untuk Mengklasifikasi Komentar Dalam Postingan

A

B Submit

C **HASIL KLASIFIKASI**

Gambar 3.12 *Input Link Page (Rancangan)*

Keterangan:

1. Label (A) *input link* berfungsi untuk memasukkan *link* pengguna untuk melakukan prediksi aspek terhadap komentar yang terdapat di dalam *link* postingan.
2. Label (B) *button submit* berfungsi untuk menjalankan proses klasifikasi.
3. Label (C) berfungsi untuk menampilkan hasil prediksi dari komentar yang terdapat dalam *link* pengguna berdasarkan aspek yang diinput.

3.4 Metode Evaluasi

Metode evaluasi dilakukan guna untuk menilai kinerja dan mengevaluasi seberapa akurat sistem berhasil dalam mengklasifikasikan komentar *cyberbullying* berbasis aspek pada metode yang digunakan. *Confusion matrix* digunakan untuk menghitung jumlah data benar dan salah. Pengukuran evaluasi dilakukan berdasarkan *confusion matrix* dengan aturan variable *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Tabel 3.9 menampilkan penerapan dari *confusion matrix*.

Tabel 3.9 *Confusion Matrix*

Clasification		Expected	
		Positive	Negative
Predicted	Positive	TP (<i>True Positive</i>)	FP (<i>False Positive</i>)
	Negative	FN (<i>False Negative</i>)	TN (<i>True Negative</i>)

Keterangan:

- **True Positive** (TP) mengacu pada sampel yang termasuk dalam *class positive* dan diklasifikasikan dengan benar.
- **True Negative** (TN) mengacu pada sampel yang termasuk dalam *class negative* dan diklasifikasikan dengan benar.
- **False Positive** (FP) merujuk pada sampel yang termasuk dalam *class negative* tetapi diklasifikasikan secara salah sebagai *class positive*.
- **False Negative** (FN) merujuk pada sampel yang termasuk dalam *class positive* tetapi diklasifikasikan secara salah sebagai *class negative*.

BAB 4

IMPLEMENTASI DAN PENGUJIAN SISTEM

4.1 Implementasi Sistem

Berdasarkan rancangan pada Bab 3 akan digunakan untuk menerapkan sistem klasifikasi berbasis aspek teks pada komentar di media sosial Instagram dengan menerapkan algoritma *Bidirectional Long-Short Term Memory* (BiLSTM). Demi mendukung pengimplementasian sistem dalam penelitian ini menggunakan perangkat keras (*hardware*) dan perangkat lunak (*software*) yang terdiri dari:

4.1.1 Spesifikasi Perangkat yang Digunakan

Perangkat keras (*hardware*) yang digunakan dalam penelitian ini dan spesifikasinya untuk mengimplementasikan sistem adalah sebagai berikut:

1. HP Notebook – 14s-cf0063TU
2. *Processor* Intel® Core™ i3-7020U (2.3 GHz, 3 MB cache, 2 cores)
3. *Memory* (RAM): 4 GB DDR4
4. *Graphic Processing Unit* (GPU): Intel® HD Graphics 620
5. *Hard Drive*: SSD 259 GB

Tidak hanya *hardware* yang digunakan, *software* juga diimplementasikan dalam penelitian ini mengimplementasikan sistem sebagai berikut:

1. Operating System: Windows 10 Home Single Language 64-bit
2. Bahasa pemrograman: *Python* versi 3.10.6, yang dilengkapi dengan sejumlah pustaka (*library*), termasuk:

<i>a. Flask</i>	<i>f. Tensorflow</i>
<i>b. Numpy</i>	<i>g. Keras</i>
<i>c. Pandas</i>	<i>h. Matplotlib</i>
<i>d. Sklearn</i>	<i>i. Seaborn</i>
<i>e. Sastrawi</i>	<i>j. Pickle</i>
3. Text Editor: *Microsoft Visual Studio Code*

4.1.2 Implementasi Rancangan Tampilan Antarmuka

Sebagaimana dirancang dalam bab sebelumnya, antarmuka (*interface*) penelitian ini akan digunakan, yang hasilnya terdiri dari:

1. Tampilan Halaman Depan (*Home Page*)

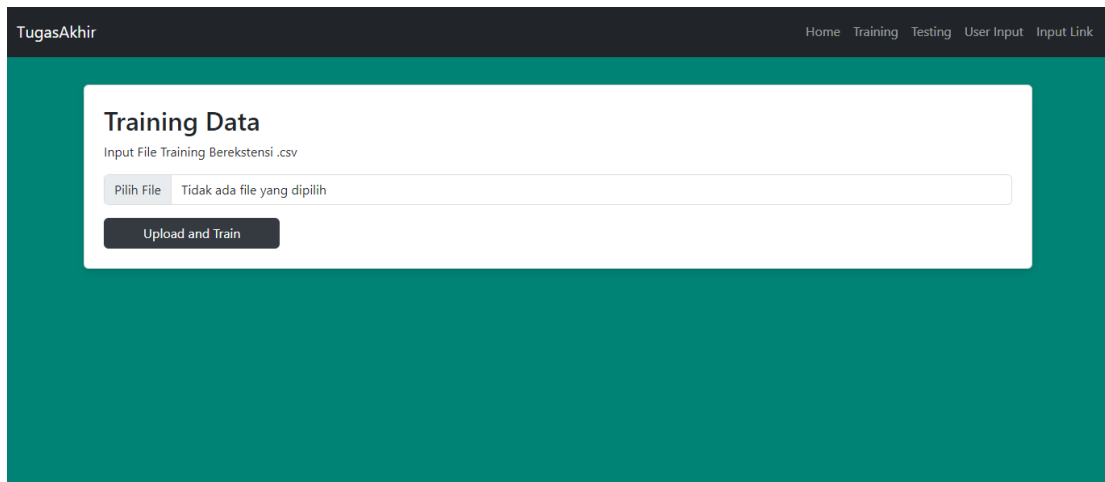
Beberapa fitur akan ditampilkan pada halaman depan (*home*) ini termasuk menu bar (*navbar*), yang berisikan menu *home*, *training*, *testing*, dan *user input* yang dapat digunakan oleh pengguna saat menjalankan semua proses tersebut. Dan pada halaman ini juga menampilkan judul penelitian, identitas penulis, navigasi ke halaman lain, dan nama instansi dari penulis akan ditampilkan. Gambar 4.1 menunjukkan rancangan *home page*.



Gambar 4.1 Home Page (Tampilan)

2. Tampilan Halaman Training (*Training Page*)

Bagian halaman *training* (*training page*) akan menampilkan form untuk pengguna yang memungkinkan pengguna untuk mengunggah file yang berekstensi *.csv* yang tersedia untuk dapat menyelesaikan proses *training* pada model. Setelah file diunggah, pengguna dapat melakukan proses *training* dengan menekan tombol “Upload and Train”. Gambar 4.2 berikut menunjukkan hasil implementasi dari bagian *training* untuk mengupload file data latih.

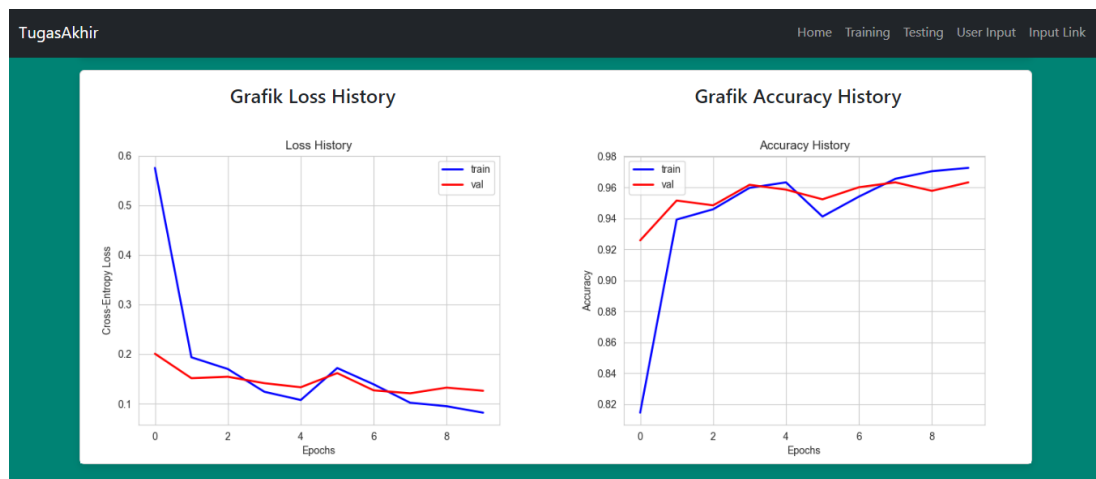


Gambar 4.2 Training Page (Tampilan)

Setelah proses *training* pada model selesai, tahapan *preprocessing* akan ditampilkan pada salah satu teks inputan. Dan juga akan menampilkan grafik nilai *accuracy* dan nilai *loss* dari model aspek. Maka pada Gambar 4.3 akan menampilkan dari tampilan tahapan *preprocessing*, dan Gambar 4.4 menunjukkan tampilan grafik dari model aspek.

	Teks Utama Asli	mukanya mirip kelinci global tv wkwk
1	Case Folding	mukanya mirip kelinci global tv wkwk
2	Pembersihan Data	mukanya mirip kelinci global tv wkwk
3	Normalisasi Data	mukanya mirip kelinci global tv wkwk
4	Stopword Removal	mukanya mirip kelinci global tv wkwk
5	Stemming	muka mirip kelinci global tv wkwk
6	Tokenisasi	[62, 474, 1571, 914, 1572, 852]

Gambar 4.3 Training Page (Tampilan Hasil Preprocessing)



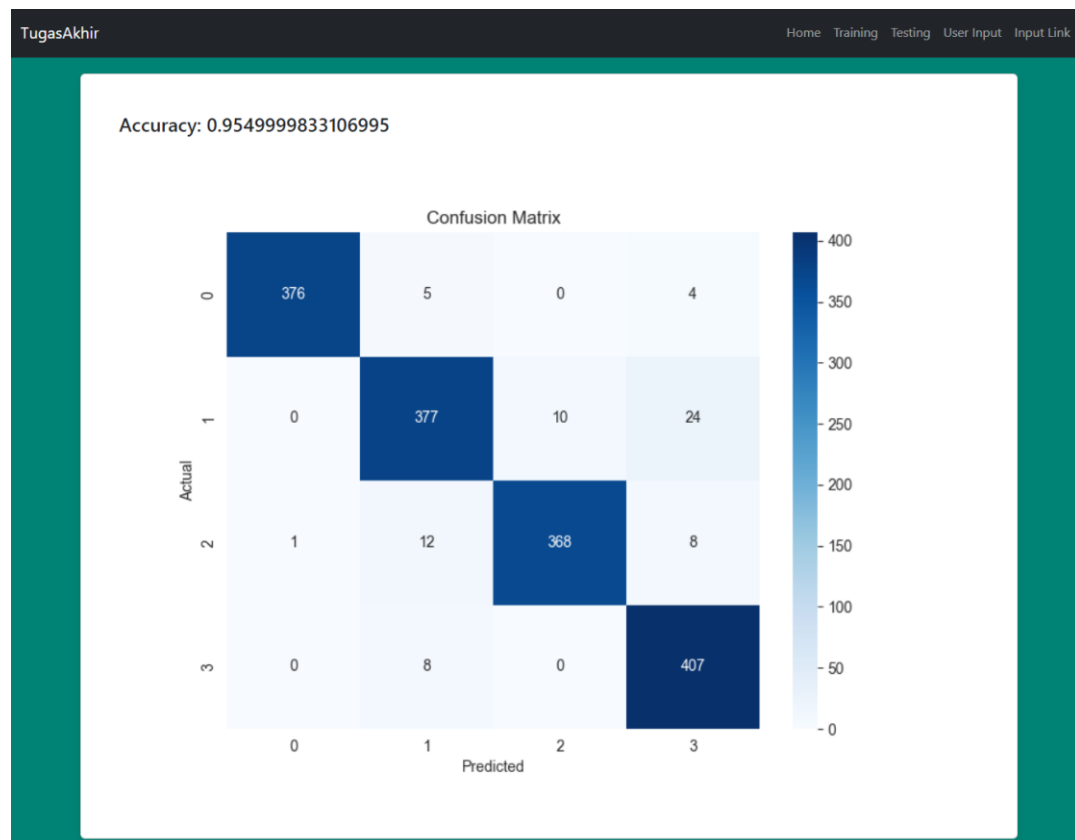
Gambar 4.4 Training Page (Tampilan Grafik Loss dan Accuracy)

3. Tampilan Halaman Testing (*Testing Page*)

Untuk memulai proses *testing* yang bertujuan untuk mengevaluasi model, maka halaman *testing* akan menampilkan form yang memungkinkan pengguna mengunggah file berekstensi *.csv*. Gambar 4.5 menampilkan bagian testing.

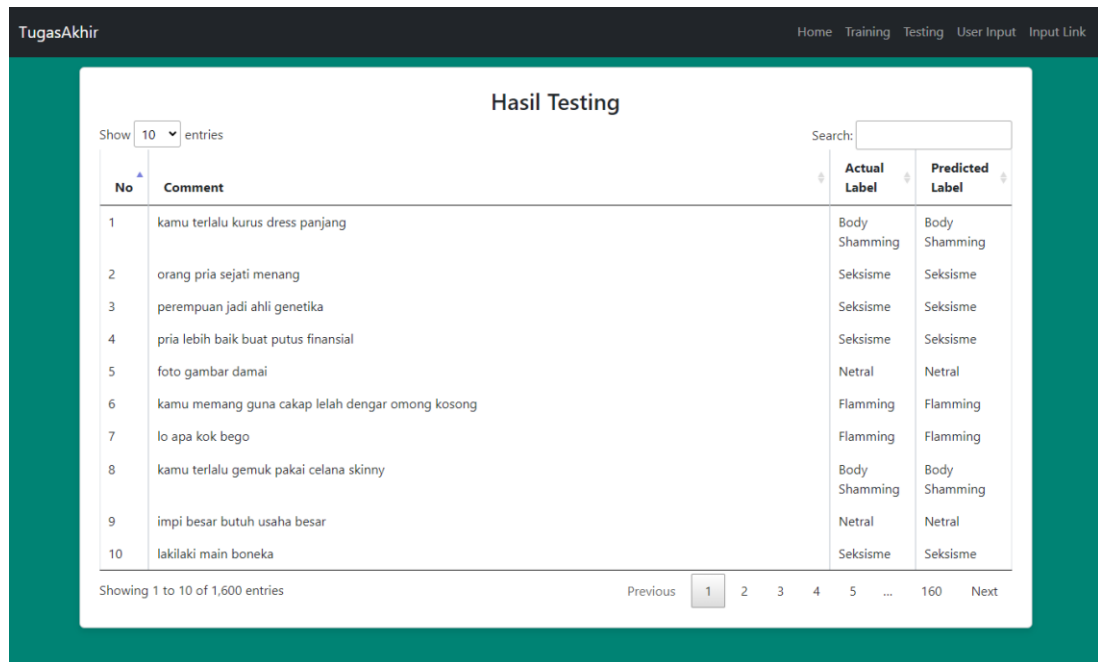
Gambar 4.5 *Testing Page (Tampilan)*

Setelah file diunggah dan proses *testing* selesai, maka hasil akan ditampilkan. Yang mana nilai akurasi prediksi model berdasarkan perbandingan antara hasil prediksi model dengan hasil sebenarnya. Gambar 4.6 berikut menampilkan hasil dari akurasi dan *confussion matrix*.



Gambar 4.6 *Tampilan Halaman Akurasi dan Confussion Matrix (Testing Page)*

Di bagian bawah juga terdapat hasil prediksi untuk setiap aspek yang berupa tabel. Seperti pada Gambar 4.7 dibawah ini.



The screenshot shows a web application interface titled 'TugasAkhir' with navigation links: Home, Training, Testing, User Input, and Input Link. The main content area is titled 'Hasil Testing'. It features a 'Show 10 entries' dropdown and a search bar. Below is a table with 10 rows of test results. The table has four columns: 'No', 'Comment', 'Actual Label', and 'Predicted Label'. The data is as follows:

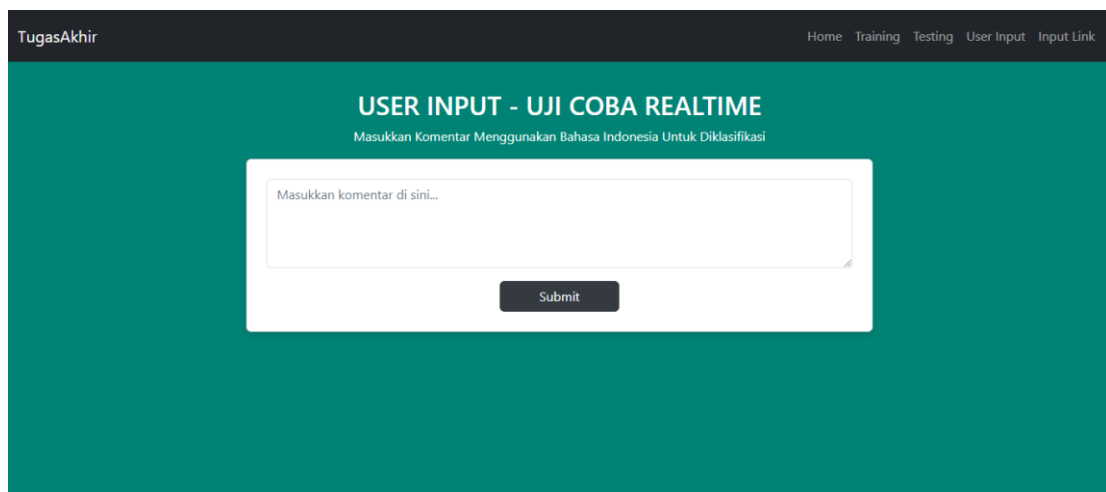
No	Comment	Actual Label	Predicted Label
1	kamu terlalu kurus dress panjang	Body Shamming	Body Shamming
2	orang pria sejati menang	Seksisme	Seksisme
3	perempuan jadi ahli genetika	Seksisme	Seksisme
4	pria lebih baik buat putus finansial	Seksisme	Seksisme
5	foto gambar damai	Netral	Netral
6	kamu memang guna cakap lelah dengar omong kosong	Flamming	Flamming
7	lo apa kok bego	Flamming	Flamming
8	kamu terlalu gemuk pakai celana skinny	Body Shamming	Body Shamming
9	impi besar butuh usaha besar	Netral	Netral
10	lakilaki main boneka	Seksisme	Seksisme

At the bottom of the table, it says 'Showing 1 to 10 of 1,600 entries'. There are pagination controls: 'Previous', '1' (selected), '2', '3', '4', '5', '...', '160', and 'Next'.

Gambar 4.7 Testing Page (Tampilan Hasil Klasifikasi)

4. Tampilan Halaman User Input (*User Input Page*)

Bagian sistem yang memungkinkan untuk proses klasifikasi dengan inputan yang terdiri dari teks komentar yang diberikan pengguna akan ditunjukkan dalam tampilan bagian user input. Halaman user input yang digunakan untuk menerima input teks pengguna dapat dilihat pada Gambar 4.8.



The screenshot shows a web application interface titled 'TugasAkhir' with navigation links: Home, Training, Testing, User Input, and Input Link. The main content area is titled 'USER INPUT - UJI COBA REALTIME'. Below the title is a subtitle: 'Masukkan Komentar Menggunakan Bahasa Indonesia Untuk Diklasifikasi'. There is a large text input field with the placeholder text 'Masukkan komentar di sini...'. Below the input field is a 'Submit' button.

Gambar 4.8 User Input Page (Tampilan)

Setelah teksnya dimasukkan oleh pengguna dan menekan tombol "Submit" maka hasil klasifikasi dari teks komentar yang dimasukkan akan muncul dan akan di

klasifikasikan berdasarkan aspek. Gambar 4.9 menunjukkan tampilan dari uji coba realtime.

Gambar 4.9 *User Input Page (Tampilan Hasil Inputan Teks)*

5. Tampilan Halaman Input Link (*Input Link Page*)

Bagian sistem yang memungkinkan untuk proses klasifikasi dengan inputan yang berupa *link* postingan yang diberikan pengguna akan ditunjukkan dalam tampilan bagian *input link*. Halaman *input link* yang digunakan untuk menerima input *link* pengguna dapat dilihat pada Gambar 4.10.

Gambar 4.10 *Input Link Page (Tampilan)*

Setelah *link* dari postingan Instagram dimasukkan oleh pengguna dan menekan tombol "Submit" maka akan menghasilkan 15 komentar teratas dari postingan dan akan diklasifikasikan setiap komentar yang akan muncul berdasarkan aspek. Gambar 4.11 menunjukkan tampilan dari inputan *link* pengguna.

TugasAkhir
Home
Training
Testing
User Input
Input Link

INPUT LINK POSTINGAN INSTAGRAM

Masukkan Link Postingan Untuk Mengklasifikasi Komentar Dalam Postingan

Hasil Klasifikasi

No	Komentar	Hasil Prediksi
1	Hamil berapa bulan arinal?	Flamming
2	Harga jagung anjlokk pak. Tolong didgr jeritan petani @jokowi @a.amran_sulaiman	Flamming
3	DILIAT LIAT GEMUK... BANYAK JUGA MAKAN DUIT HARAM 🤔🤔	Body Shamming
4	Dikit lg perutnya meledak	Body Shamming
5	Lampung barat, sekolah kopi	Flamming
6	Perutnya buncit kebanyakan mkn.....?hayo silahkan di isi	Body Shamming
7	Pak gub tolong kami..kami mau mengurus naik pangkat di kantor BKD..tapi kami trauma..takut di pukuli kaya anak IPDN..apamng sekarang kantor BKD provinsi sarang penjahat y	Netral
8	diet ndut, dasar gendut	Body Shamming
9	Tai lo	Flamming
10	Goblok sia	Flamming
11	@official.kpk mohon ditindak lanjuti	Netral
12	BENTAR LAGI MELEDAK TUH PERUT 🤔	Body Shamming
13	PERUTNYA ISINYA APASI WOUYYY, UANG HARAM !	Body Shamming
14	Cuak	Flamming
15	Gendut gemesin yaa 🤔	Flamming

Gambar 4.11 *Input Link (Tampilan Hasil dari Inputan Link)*

4.2 Implementasi Model

Dalam pembahasan ini, penulis akan membahas pelaksanaan dan penerapan algoritma untuk mencapai hasil model yang terbaik. Algoritma yang diterapkan adalah *Bidirectional Long-Short Term Memory* (BiLSTM) berfungsi untuk mewujudkan sistem. Adapun implementasi model yang akan diterapkan yaitu model *training* dan model *testing* yang dijelaskan berikut ini:

4.2.1 Model Training

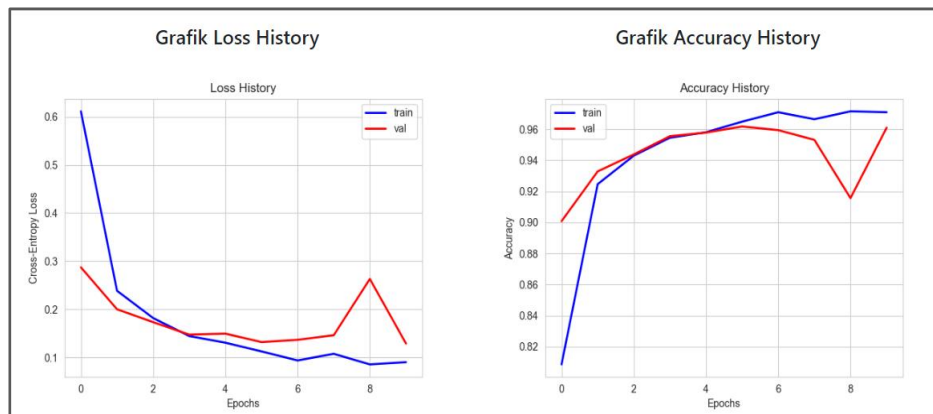
Untuk menghasilkan model yang terbaik dan dengan akurasi tertinggi dalam melakukan klasifikasi komentar, penelitian ini menerapkan algoritma *Bidirectional Long-Short Term Memory* (BiLSTM), yang digunakan melalui percobaan pelatihan untuk mengganti beberapa *hyperparameter*. Adanya lapisan *dropout* dan fungsi dari *Early Stopping* pada model dapat membantu mencegah terjadinya *overfitting*. Tabel 4.1 berikut

menunjukkan beberapa percobaan yang dilakukan penulis dalam proses pengubahan *hyperparameter*.

Tabel 4.1 Percobaan Pengubahan Hyperparameter

Unit LSTM	Epoch	Accuracy	Validation Loss
32	5	0.9625	0.1239
64	10	0.9638	0.3057
32	15	0.9669	0.1729
64	20	0.9651	0.2491
32	25	0.9619	0.2269
64	30	0.9656	0.2825

Dalam penelitian ini, penulis mengidentifikasi bahwa model dengan kinerja yang optimal diperoleh menggunakan ukuran *batch* sebesar 32, dengan *epoch* 15. Untuk proses pelatihan model yang menghasilkan grafik *loss history* dan grafik *accuracy history* pada Gambar 4.12. Grafik ini menampilkan visualisasi yang komprehensif mengenai bagaimana model belajar dan menyempurnakan kemampuannya seiring waktu.



Gambar 4.12 Grafik Loss dan Accuracy

Pada Gambar 4.12, grafik akurasi untuk data *training* menunjukkan bahwa model mencapai tingkat *accuracy* mencapai 0.9704 serta nilai *loss* mencapai 0.1140. Sementara itu, data validasi menunjukkan tingkat *accuracy* mencapai angka 0.9669 dengan nilai *loss* mencapai 0.1729. Dengan hasil yang ditunjukkan bahwa model memiliki kinerja yang cukup baik dalam memprediksi data baru, menunjukkan generalisasi yang memadai dari data pelatihan dan data validasi.

4.2.2 Model Testing

Dalam tahap pengujian sistem, model yang telah dikembangkan akan disimpan setelah proses *training* dan diuji menggunakan data *testing* untuk mengevaluasi kemampuan sistem dalam mengklasifikasikan komentar Instagram dengan metode *Bidirectional Long-Short Term Memory* (BiLSTM). Pengujian ini memanfaatkan 1600 data baru yang belum dikenali oleh model. Hasil pengujian model disajikan Tabel 4.2.

Tabel 4.2 Hasil Pengujian Sistem

Comment	Actual Label	Predicted Label
kamu terlalu kurus dress panjang	Body Shamming	Body Shamming
orang pria sejati menang	Seksisme	Seksisme
perempuan jadi ahli genetika	Seksisme	Seksisme
pria lebih baik buat putus finansial	Seksisme	Seksisme
foto gambar damai	Netral	Netral
kamu memang guna cakap lelah dengar omong kosong	Flamming	Flamming
kamu terlalu gemuk pakai celana skinny	Body Shamming	Body Shamming
percuma aja ngomong kalau cuma bikin orang jadi tambah bingung	Flamming	Flamming
perut terlalu besar coba kurang konsumsi gula	Body Shamming	Body Shamming
wanita lebih utama tampil fisik prestasi akademis	Seksisme	Seksisme
.....
kamu sangat fotogenik	Netral	Netral
lo bikin semua jadi kacau	Flamming	Flamming
wanita lebih baik kerja butuh sentuh lembut	Seksisme	Seksisme
outfitmu buat lihat sangat elegan	Netral	Netral

pria lebih unggul rencana uang	Seksisme	Seksisme
pernah ngertiin otak tuh kayak kapas	Flamming	Flamming
tubuh terlalu besar dress cocktail	Body Shamming	Body Shamming
kamu memang punya otak ngomongnya aja aneh	Flamming	Flamming
rezeki gaakan kemanamana mas robby semangat	Netral	Netral

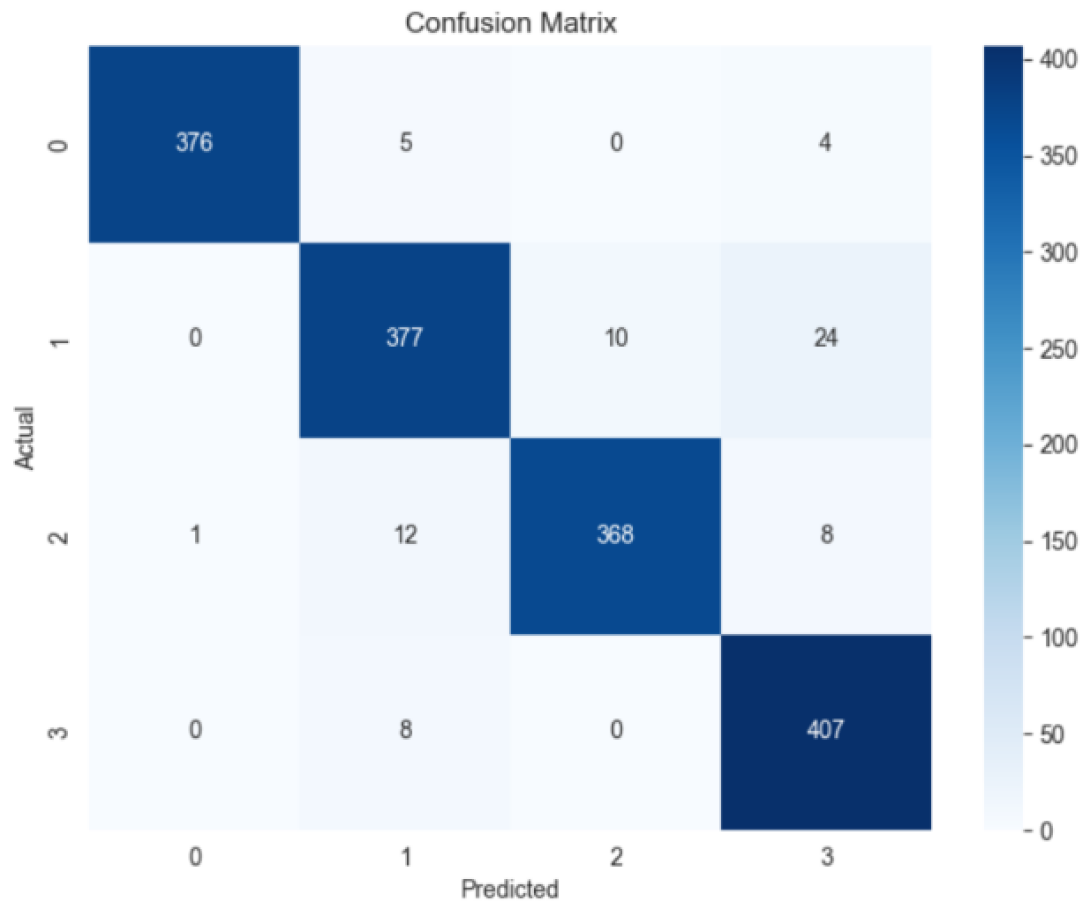
4.3 Metode Evaluasi

Pada tahap evaluasi model, perhitungan nilai akurasi dilakukan menggunakan matriks akurasi. Oleh karena itu, diperlukan data yang mencakup hasil prediksi benar dan salah oleh model. Dalam penelitian ini, dataset pengujian yang digunakan terdiri dari 1600 sampel yang berasal dari kumpulan data dengan *class* tertentu. *Class* data pada Tabel 4.3:

Tabel 4.3 *Class Dataset*

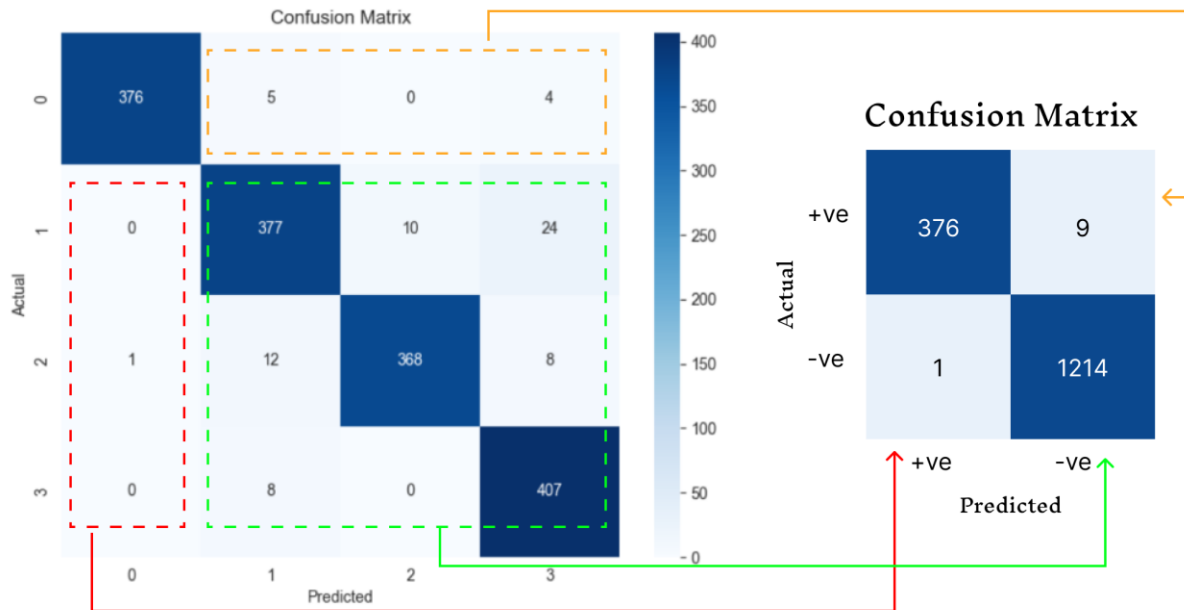
Class	N ^o of Samples
0 (Sexism)	385
1 (Flaming)	411
2 (Body-Shaming)	389
3 (Netral)	415

Sistem evaluasi yang diterapkan dalam penelitian ini adalah *confusion matrix*. *Confusion matrix* diperoleh dengan melatih pengklasifikasi dan mengevaluasi model yang telah dilatih pada set pengujian. Hasil *confusion matrix* pada Gambar 4.13 di bawah ini.



Gambar 4.13 *Confusion Matrix (Semua Class)*

Confusion matrix ini dapat diubah menjadi matriks jenis *one-vs-all* (*binary-class confusion matrix*) untuk menghitung matriks berdasarkan kelas seperti *accuracy*, *precision*, *recall*, dan lain-lain. Proses konversi matriks menjadi matriks *one-vs-all* untuk data kelas 0 akan terlihat dalam Gambar 4.14 ini.



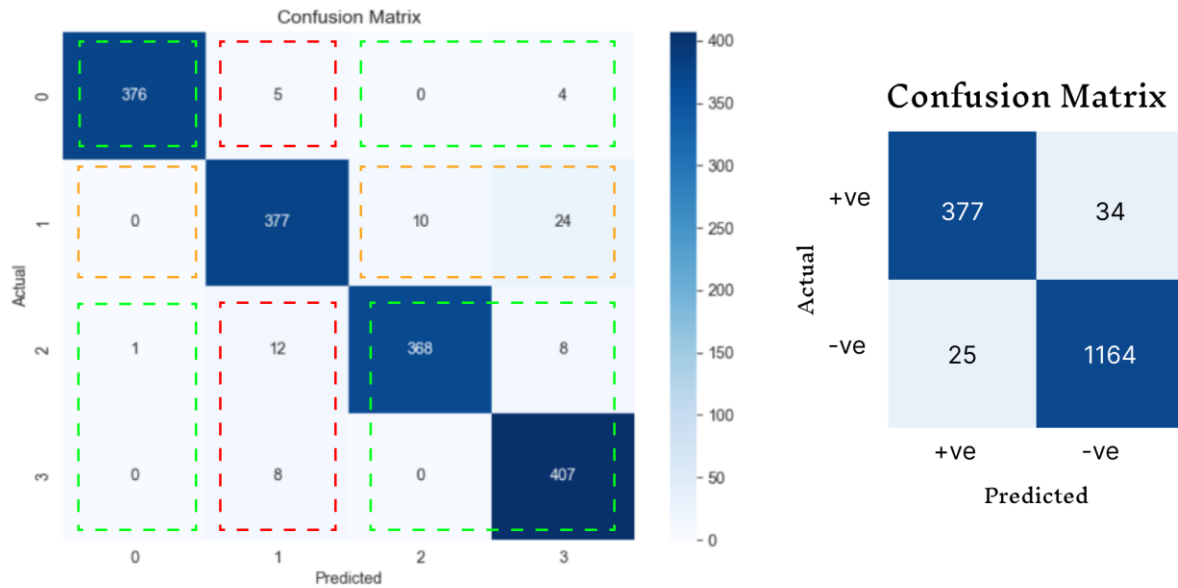
Gambar 4.14 Konversi Matriks ke Class 0 (Sexism)

Pada Gambar 4.14, yang merupakan hasil konversi matriks untuk kelas seksisme, perhitungan menunjukkan bahwa terdapat 376 data yang diprediksi dengan benar sebagai kelas seksisme. Selain itu, terdapat 1 data yang diprediksi dengan benar namun tidak termasuk dalam kelas seksisme. Terdapat pula 9 data yang diprediksi secara salah sebagai kelas seksisme dan 1214 data yang diprediksi secara salah sebagai bukan kelas seksisme. Informasi lebih rinci mengenai hasil analisis *confusion matrix class sexism* Tabel 4.4.

Tabel 4.4 Keterangan *Confusion Matrix Class Sexism*

No.	Class Sexism	Jumlah
1.	TP (True Positive)	376
2.	FP (False Positive)	1
3.	FN (False Negative)	9
4.	TN (True Negative)	1214

Demikian pula untuk class-1, *confusion matrix one-vs-all* yang dikonversi akan terlihat seperti Gambar 4.15:



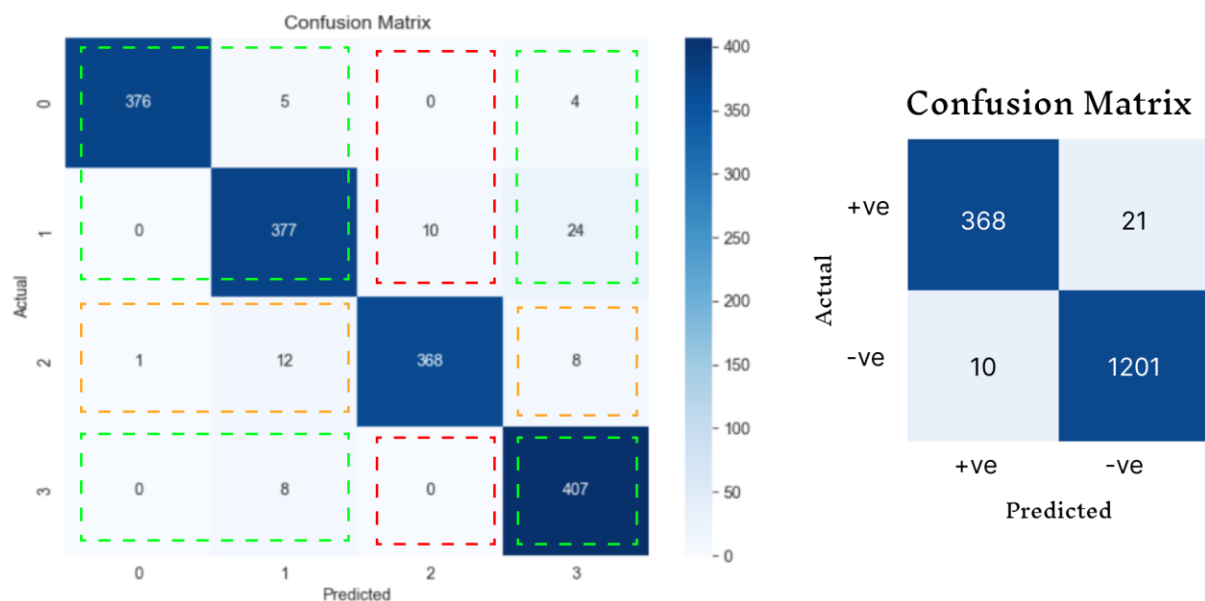
Gambar 4.15 Konversi Matriks ke Class 1 (Flaming)

Gambar 4.15, yang merupakan hasil konversi matriks untuk kelas *flaming*, perhitungan menunjukkan bahwa terdapat 377 data yang diprediksi dengan benar sebagai kelas *flaming*. Selain itu, terdapat 25 data yang diprediksi dengan benar namun tidak termasuk dalam kelas *flaming*. Terdapat pula 34 data yang diprediksi secara salah sebagai kelas *flaming* dan 1164 data yang diprediksi secara salah sebagai bukan kelas *flaming*. Informasi yang lebih terperinci mengenai hasil analisis *confusion matrix class flaming* pada Tabel 4.5.

Tabel 4.5 Keterangan Confusion Matrix Class Flaming

No.	Class Flaming	Jumlah
1.	TP (True Positive)	377
2.	FP (False Positive)	25
3.	FN (False Negative)	34
4.	TN (True Negative)	1164

Demikian pula untuk class-2, *confussion matrix one-vs-all* yang dikonversi terlihat pada Gambar 4.16:



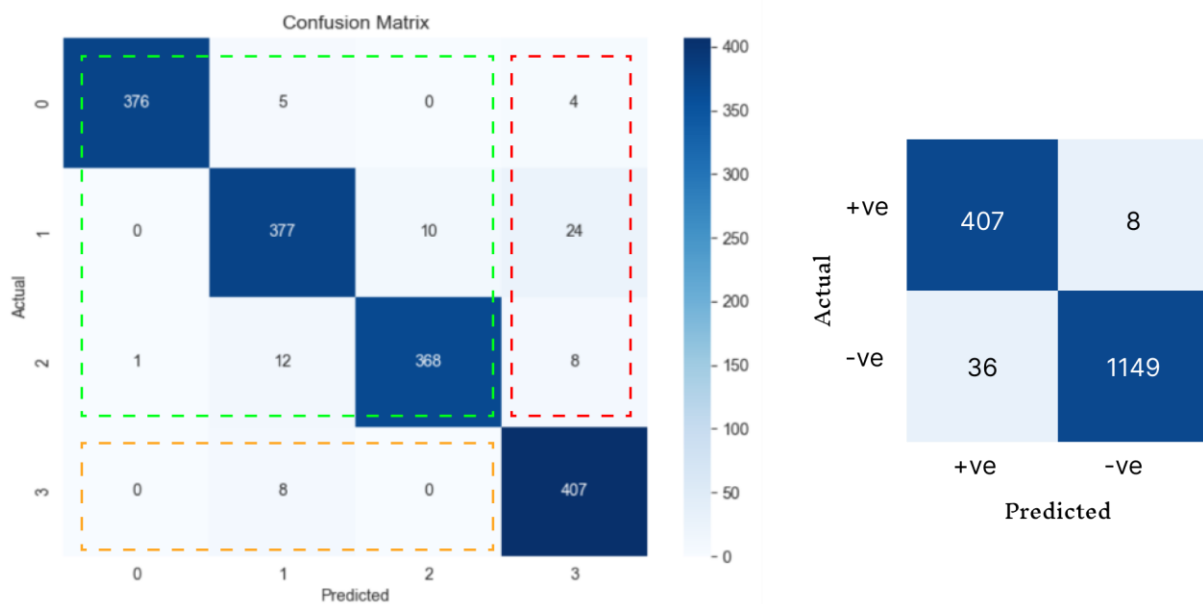
Gambar 4.16 Konversi Matriks ke Class 2 (*Body Shaming*)

Pada Gambar 4.16, yang menampilkan hasil konversi matriks untuk kelas *body-shaming*, analisis menunjukkan bahwa terdapat 368 data yang berhasil diprediksi dengan benar sebagai *body-shaming*. Selain itu, ada 10 data yang diprediksi dengan benar tetapi sebenarnya tidak termasuk dalam kelas *body-shaming*. Sementara itu, terdapat 21 data yang diprediksi secara salah sebagai *body-shaming*, dan 1201 data yang diprediksi secara salah sebagai bukan *body-shaming*. Informasi lebih rinci mengenai kelas *body-shaming* yang dihasilkan dari *confusion matrix* dapat ditemukan Tabel 4.6.

Tabel 4.6 Keterangan *Confusion Matrix Class Body Shaming*

No.	Class Body Shaming	Jumlah
1.	TP (True Positive)	368
2.	FP (False Positive)	10
3.	FN (False Negative)	21
4.	TN (True Negative)	1201

Demikian pula untuk class-3, *confussion matrix one-vs-all* yang dikonversi pada Gambar 4.17:



Gambar 4.17 Konversi Matriks ke Class 3 (Netral)

Pada Gambar 4.17, yang merupakan hasil konversi matriks untuk kelas *netral*, perhitungan menunjukkan bahwa terdapat 407 data yang diprediksi dengan benar sebagai kelas *netral*. Selain itu, terdapat 36 data yang diprediksi dengan benar tetapi sebenarnya tidak termasuk dalam kelas *netral*. Selanjutnya, terdapat 8 data yang diprediksi secara salah sebagai kelas *netral*, dan 1149 data yang diprediksi secara salah sebagai bukan kelas *netral*. Informasi lebih terperinci mengenai kelas *netral* yang dihasilkan dari *confusion matrix* dapat dilihat Tabel 4.7.

Tabel 4.7 Keterangan Confusion Matrix Class Netral

No.	Class Netral	Jumlah
1.	TP (True Positive)	407
2.	FP (False Positive)	36
3.	FN (False Negative)	8
4.	TN (True Negative)	1149

Dengan menerapkan konsep ini, akan mempermudah untuk menghitung matriks seperti *accuracy*, *precision*, *recall* dan *f1-scores* untuk setiap kelas, kemudia dapat mentabulasi hasilnya untuk analisis lebih lanjut.

- Kalkulasi *Precision, Recall, F1-Score Class Sexism*:

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{376}{376 + 1} = \frac{376}{377} = 0.9973$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{376}{376 + 9} = \frac{376}{385} = 0.9766$$

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0.9973 \times 0.9766}{0.9973 + 0.9766} = \frac{1.9479}{1.9739} = 0.9868$$

- Kalkulasi *Precision, Recall, F1-Score Class Flaming*:

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{377}{377 + 25} = \frac{377}{402} = 0.9378$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{377}{377 + 34} = \frac{377}{411} = 0.9173$$

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0.9378 \times 0.9173}{0.9378 + 0.9173} = \frac{1.7205}{1.8551} = 0.9274$$

- Kalkulasi *Precision, Recall, F1-Score Class Body Shaming*:

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{368}{368 + 10} = \frac{368}{378} = 0.9735$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{368}{368 + 21} = \frac{368}{389} = 0.9460$$

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0.9735 \times 0.9460}{0.9735 + 0.9460} = \frac{1.8418}{1.9195} = 0.9595$$

- Kalkulasi *Precision, Recall, F1-Score Class Netral*:

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{407}{407 + 36} = \frac{407}{443} = 0.9187$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{407}{407 + 8} = \frac{407}{415} = 0.9807$$

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0.9187 \times 0.9807}{0.9187 + 0.9807} = \frac{1.8019}{1.8994} = 0.9487$$

Hasil dari perhitungan *precision, recall*, dan *f1-score* untuk setiap class pada Tabel 4.8.

Tabel 4.8 *Perhitungan Kalkulasi*

Class	Precision (%)	Recall (%)	F1-Score (%)
0 (Sexism)	99.73%	97.66%	98.68%
1 (Flaming)	93.78%	91.73%	92.74%
2 (Body Shaming)	97.35%	94.60%	95.95%
3 (Netral)	91.87%	98.07%	94.87%

F1-Score yang dirata-ratakan secara mikro merupakan matriks global yang dihitung dengan mempertimbangkan *True Positive* (TP) secara keseluruhan, yaitu jumlah TP berdasarkan setiap *class* (dari matriks *one-vs-all*), serta *False Positive* (FP) dan *False Negative* (FN) secara keseluruhan. Hasilnya seperti berikut:

- Net TP (*True Positive*) = $376 + 377 + 368 + 407 = 1528$
- Net FP (*False Positive*) = $(0 + 1 + 0) + (5 + 12 + 8) + (0 + 10 + 0) + (4 + 24 + 8) = 72$
- Net FN (*False Negative*) = $(5 + 0 + 4) + (0 + 10 + 24) + (1 + 12 + 8) + (0 + 8 + 0) = 72$

Jika diperhatikan untuk setiap *confusion matrix* nilai net FP dan net FN akan memiliki nilai sama. Oleh karena itu, *micro precision* dan *micro recall* dapat dihitung seperti berikut:

- Micro Precision = $\frac{Net\ TP}{Net\ TP + Net\ FP} = \frac{1528}{1528 + 72} = \frac{1528}{1600} = 0.955 = 95.5\%$
- Micro Recall = $\frac{Net\ TP}{Net\ TP + Net\ FN} = \frac{1528}{1528 + 72} = \frac{1528}{1600} = 0.955 = 95.5\%$
- Micro F1-Score = $\frac{2 \times Micro\ Precision \times Micro\ Recall}{Micro\ Precision + Micro\ Recall}$
 $= \frac{2 \times 0.955 \times 0.955}{0.955 + 0.955}$
 $= \frac{1.8241}{1.91}$
 $= 0.955 = 95.5\%$

Dengan nilai rata-rata *weight* yang diperoleh dengan menghitung nilai rata-rata *weight* dari sampel nilai-nilai kelas yang didapat. Dengan demikian, nilai *weight* yang dihasilkan adalah:

- Weighted Precision = $\frac{(99.73 \times 385) + (93.78 \times 411) + (97.35 \times 389) + (91.87 \times 415)}{377 + 402 + 378 + 443} \%$
 $= \frac{152.934,83}{1600} \%$
 $= 95.58\%$
- Weighted Recall = $\frac{(97.66 \times 385) + (91.73 \times 411) + (94.60 \times 389) + (98.07 \times 415)}{377 + 402 + 378 + 443}$
 $= \frac{152.798,58}{1600} \%$
 $= 95.5\%$
- Weighted F1-Score = $\frac{(98.68 \times 385) + (92.74 \times 411) + (95.95 \times 389) + (94.87 \times 415)}{377 + 402 + 378 + 443}$
 $= \frac{152.803,54}{1600} \%$
 $= 95,50\%$

Kurva *Receiver Operating Characteristic* (ROC) merupakan representasi dari grafis yang memplot ‘*true positive rate*’ terhadap ‘*false positive rate*’ pada berbagai pengaturan ambang batas. Meskipun kurva ROC umumnya digunakan untuk model klasifikasi biner, kurva ini juga dapat diadaptasikan untuk pengaturan multikelas.

True Positive Rate (TPR) didefinisikan secara identik dengan sensitivitas (*recall*), yang mengukur proporsi sampel dalam kelas positif dari suatu dataset yang diklasifikasikan dengan benar oleh model prediktif. Oleh karena itu, rumus untuk menghitung TPR adalah sebagai berikut:

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

Dari hasil pengujian klasifikasi komentar *cyberbullying* pada media sosial Instagram menggunakan algoritma *Bidirectional Long-Short Term Memory* (BiLSTM), diperoleh tingkat akurasi sebesar 95.5%. Hasil ini menunjukkan bahwa model *Bidirectional Long-Short Term Memory* (BiLSTM) mampu dalam mengklasifikasikan komentar *cyberbullying* berdasarkan aspek-aspeknya di media sosial Instagram dengan tingkat keakuratan yang cukup tinggi.

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penerapan algoritma *Bidirectional Long-Short Term Memory* (BiLSTM) dalam mengklasifikasikan komentar *cyberbullying* di media sosial Instagram berdasarkan aspeknya, beberapa kesimpulan dapat ditarik yaitu:

1. Dengan membagi data latih dan data uji dalam proporsi 80% dan 20%, ditemukan bahwa model *Bidirectional Long-Short Term Memory* (BiLSTM) yang dibangun mencapai rata-rata akurasi aspek sebesar 95.5%.
2. Berdasarkan tingkat akurasi yang dicapai, ketepatan model *Bidirectional Long-Short Term Memory* yang telah dikembangkan menunjukkan hasil yang baik untuk digunakan dalam klasifikasi komentar *cyberbullying* pada media sosial Instagram berbahasa Indonesia. Model ini mampu dalam mengklasifikasikan komentar ke dalam empat aspek, yaitu *sexism*, *flaming*, *body shaming*, dan netral.

5.2 Saran

Penulis memberikan beberapa saran untuk pengembangan sistem dalam penelitian mendatang, yaitu sebagai berikut:

1. Disamping mengklasifikasikan komentar *cyberbullying* dalam bahasa Indonesia, diperlukan penelitian lanjutan yang dapat memproses komentar dalam bahasa Inggris. Sistem juga sebaiknya diperluas untuk dapat memahami tanda baca serta emotikon, sehingga analisis tidak hanya terbatas pada teks saja.
2. Penelitian ini memfokuskan pada klasifikasi *cyberbullying* di media sosial Instagram. Untuk penelitian berikutnya, disarankan untuk mengeksplorasi platform lain seperti TikTok, YouTube, Facebook dan platform lainnya guna memperluas ruang lingkup studi ini.
3. Pada penelitian berikutnya, disarankan untuk menambahkan dataset lebih beragam dan lebih banyak dari penelitian ini.

DAFTAR PUSTAKA

- Agrawal, S., & Awekar, A. (2018, March). Deep learning for detecting cyberbullying across multiple social media platforms. In *European conference on information retrieval* (pp. 141-153). Cham: Springer International Publishing.
- Ahmed, S., Saif, A. S., Hanif, M. I., Shakil, M. M. N., Jaman, M. M., Haque, M. M. U., ... & Sabbir, H. M. (2021). Att-BiL-SL: Attention-based Bi-LSTM and sequential LSTM for describing video in the textual formation. *Applied sciences*, 12(1), 317.
- Ariesto, A. (2009). Pelaksanaan Program Anti-Bullying Teacher Empowerment Program di Sekolah. Skripsi. Fakultas Ilmu Sosial dan Ilmu Politik, Universitas Indonesia.
- Atmoko, B. D. (2012). Instagram Handbook Tips Fotografi Ponsel. Jakarta: Mediakita.
- Ayres, M. M., Friedman, C. K., & Leaper, C. (2009). Individual and situational factors related to young women's likelihood of confronting sexism in their everyday lives. *Sex Roles*, 61(7), 449-460.
- Bauman, S. (2014). *Cyberbullying: What counselors need to know*. John Wiley & Sons.
- Brownlee, J. (2017). How to use word embedding layers for deep learning with Keras. Retrieved from *Machine Learning Mastery website*: <https://machinelearningmastery.com/use-word-embedding-layers-deeplearning-keras>.
- Candra, R. M., & Rozana, A. N. (2020). Klasifikasi Komentar Bullying pada Instagram Menggunakan Metode K-Nearest Neighbor. *IT Journal Research and Development*, 5(1), 45-52.
- Chamidah, N., & Sahawaly, R. (2021). Comparison Support Vector Machine and Naive Bayes Methods for Classifying Cyberbullying in Twitter. *J. Ilm. Tek. Elektro Komput. dan Inform*, 7(2), 338.
- Ciampaglia, G. L. (2017). Can online gaming ditch its sexist ways?. *theconversation.com*. (Online). <https://theconversation.com/can-online-gaming-ditch-its-sexist-ways-74493> (accessed May. 20, 2024).
- Coloroso, B. (2007). Stop bullying (memutus rantai kekerasan anak dari prasekolah hingga SMU). Jakarta: Ikrar Mandiri Abadi.

- Cui, Z., Ke, R., Pu, Z., & Wang, Y. (2018). Deep bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction. *arXiv preprint arXiv:1801.02143*.
- Dai, J., & Liu, X. (2014). Approach for text classification based on the similarity measurement between normal cloud models. *The Scientific World Journal*, 2014.
- Damanik, G. N., & Djuwita, R. (2019). Gambaran perundungan pada siswa tingkat SMA di Indonesia. *Jurnal Psikogenesis*, 7(1), 28-40.
- Eisenberg, M. E., Neumark-Sztainer, D., & Story, M. (2003). Associations of weight-based teasing and emotional well-being among adolescents. *Archives of pediatrics & adolescent medicine*, 157(8), 733-738.
- Endah, T. (2016). Klasifikasi Cyber Bullying Pada Media Sosial Twitter Dengan Menggunakan Algoritma Naïve Bayes. *Institut Teknologi Sepuluh Nopember*.
- Fadli, H. F., & Hidayatullah, A. F. (2021). Identifikasi cyberbullying pada media sosial twitter menggunakan metode lstm dan bilstm. *AUTOMATA*, 2(1).
- Felmlee, D., Inara Rodis, P., & Zhang, A. (2020). Sexist slurs: Reinforcing feminine stereotypes online. *Sex roles*, 83(1), 16-28.
- Fimela. (2018). KPAI: Bullying Pada Anak Di Media Sosial Semakin Meningkat Tiap Tahunnya. *fimela.com*. (Online). <https://www.fimela.com/parenting/read/3807829/kpai-bullying-pada-anak-di-media-sosial-semakin-meningkat-tiap-tahunnya> (accessed Jun. 27, 2023).
- Frison, E., & Eggermont, S. (2017). Browsing, posting, and liking on Instagram: The reciprocal relationships between different types of Instagram use and adolescents' depressed mood. *Cyberpsychology, Behavior, and Social Networking*, 20(10), 603-609.
- Ganta, R. (2021). Word Embedding and Word2Vec. *raviteja-ganta.github.io*. (Online). <https://raviteja-ganta.github.io/Word-Embeddings> (accessed July. 25, 2023).
- Geitgey, A. (2018). Natural Language Processing is Fun!. *medium.com*. (Online). <https://medium.com/@ageitgey/natural-language-processing-is-fun-9a0bff37854e> (accessed July. 25, 2023).
- Graves, A., Jaitly, N., & Mohamed, A. R. (2013, December). Hybrid speech recognition with deep bidirectional LSTM. In *2013 IEEE workshop on automatic speech recognition and understanding* (pp. 273-278). IEEE.

- Gurusamy, V., & Kannan, S. (2014, October). Preprocessing techniques for text mining. In *Conference Paper RTRICS, India*.
- Hermawan, L., & Ismiati, M. B. (2020). Pembelajaran text preprocessing berbasis simulator untuk mata kuliah information retrieval. *Jurnal Transformatika*, 17(2), 188-199.
- Hilmawan, M. D. (2022). Deteksi Sarkasme Pada Judul Berita Berbahasa Inggris Menggunakan Algoritme Bidirectional LSTM. *Journal of Dinda: Data Science, Information Technology, and Data Analytics*, 2(1), 46-51.
- Ismiati, M. B. (2018). Deteksi Komentar Negatif Di Instagram Menggunakan Algoritma *Naïve Bayes Classifier*. *Universitas Katolik Musi Charitas*.
- Johnson, K. (2014). *Overt and Inferential Sexist Language in the Video Game Industry* (Doctoral dissertation, University of Oregon).
- Kannan, S., Gurusamy, V., Vijayarani, S., Ilamathi, J., Nithya, M., Kannan, S., & Gurusamy, V. (2014). Preprocessing techniques for text mining. *International Journal of Computer Science & Communication Networks*, 5(1), 7-16.
- Karyanti, M. P., & Aminudin, S. P. (2019). *Cyberbullying & Body Shaming*. Penerbit K-Media.
- Katadata.co.id. (2017). 2022, Pengguna Media Sosial Indonesia Mencapai 125 Juta. *databoks.katadata.co.id*. (Online).
<https://databoks.katadata.co.id/datapublish/2017/08/22/2022-pengguna-media-sosial-indonesia-mencapai-125-juta> (accessed July. 20, 2023).
- Katadata.co.id. (2019). Survei APJII: 49% Pengguna Internet Pernah Dirisak di Medsos. *databoks.katadata.co.id*. (Online).
<https://databoks.katadata.co.id/datapublish/2019/05/16/survei-apjii-49-pengguna-internet-pernah-dirisak-di-medsos> (accessed July. 20, 2023).
- Kemenpppa. (2016). Januari – Ratas Bullying KPP-PA. *kemenpppa.go.id*. (Online).
<https://www.kemenpppa.go.id/index.php/page/read/32/952/januari-ratas-bullying-kpp-pa> (accessed July. 21, 2023).

- Kominfo. (2015). “Kemkominfo: Internet Jadi Referensi Utama Mengakses Berita dan Informasi”. *kominfo.go.id*. (Online).
[\(https://www.kominfo.go.id/index.php/content/detail/5421/Kemkominfo%3A+Internet+Jadi+Referensi+Utama+Mengakses+Berita+dan+Informasi/0/berita_satker#:~:text=Jakarta%2C%20Kominfo%20%2D%20Social%20media%20adalah,%20dan%20kolaborasi%20\(collaborating\)](https://www.kominfo.go.id/index.php/content/detail/5421/Kemkominfo%3A+Internet+Jadi+Referensi+Utama+Mengakses+Berita+dan+Informasi/0/berita_satker#:~:text=Jakarta%2C%20Kominfo%20%2D%20Social%20media%20adalah,%20dan%20kolaborasi%20(collaborating)) (accessed Mar. 11, 2023).
- Kominfo. (2013). “Kominfo: Pengguna Internet di Indonesia 63 Juta Orang”. *kominfo.go.id.kominfo.go.id*. (Online).
[\(https://kominfo.go.id/index.php/content/detail/3415/Kominfo+%3A+Pengguna+Internet+di+Indonesia+63+Juta+Orang/0/berita_satker\)](https://kominfo.go.id/index.php/content/detail/3415/Kominfo+%3A+Pengguna+Internet+di+Indonesia+63+Juta+Orang/0/berita_satker) (accessed Mar. 26, 2023).
- Kurniawan, M., Kusriani, K., & Arief, M. R. (2021). Part of speech tagging pada teks Bahasa Indonesia dengan BiLSTM+ CNN+ CRF dan ELMo. *Jurnal Eksplorasi Informatika*, 11(1), 29-37.
- Le, X. H., Ho, H. V., Lee, G., & Jung, S. (2019). Application of long short-term memory (LSTM) neural network for flood forecasting. *Water*, 11(7), 1387.
- Liddy, E. D. (2001). *Natural language processing*. Encyclopedia of Library and Information Science. New York.
- Lidwina, A. (2019). Motivasi Utama Orang Indonesia Bermedsos Ternyata Bukan Cari Teman. *databoks.katadata.co.id*. (Online).
[\(https://databoks.katadata.co.id/datapublish/2019/12/27/inilah-motivasi-orang-indonesia-menggunakan-media-sosial\)](https://databoks.katadata.co.id/datapublish/2019/12/27/inilah-motivasi-orang-indonesia-menggunakan-media-sosial) (accessed Mar. 26, 2023).
- Liu, C., Liu, Z., & Yuan, G. (2020). The longitudinal influence of cyberbullying victimization on depression and posttraumatic stress symptoms: The mediation role of rumination. *Archives of Psychiatric Nursing*, 34(4), 206-210.
- Lu, P., Xi, H., Zhou, B., Zhang, H., Lin, Y., Chen, L., ... & Hu, Y. (2020). A New Multichannel Parallel Network Framework for the Special Structure of Multilead ECG. *Journal of Healthcare Engineering*, 2020.
- Luqyana, W. A., Cholissodin, I., & Perdana, R. S. (2018). Analisis Sentimen Cyberbullying Pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(11), 4704-4713.

- Mahdi, M. I. 2022. “Pengguna Media Sosial di Indonesia Capai 191 Juta pada 2022”. *dataindonesia.id*. <https://dataindonesia.id/digital/detail/pengguna-media-sosial-di-indonesia-capai-191-juta-pada-2022> (accessed Mar. 10, 2023).
- MicrosoftResearch. (2021). Studi Terbaru dari Microsoft Menunjukkan Peningkatan Digital Civility (Keadaban Digital) di Seluruh Kawasan Asia-Pacific Selama Masa Pandemi. *news.microsoft.com*. (Online). <https://news.microsoft.com/id-id/2021/02/11/studi-terbaru-dari-microsoft-menunjukkan-peningkatan-digital-civility-keadaban-digital-di-seluruh-kawasan-asia-pacific-selama-masa-pandemi/> (accessed Mar. 10, 2023).
- Monkeylearn. (2023). Text Classofocation: What it is And Why it Matters. *monkeylearn.com*. (Online). <https://monkeylearn.com/text-classification/> (accessed Jul. 23, 2023).
- Mulyawan, R. (2023). Memahami Pengertian Embedding (Penyematan): Manfaat, Cara Melakukan dan Perbedaannya dengan Uploading!. *rifqimulyawan.com*. (Online). <https://rifqimulyawan.com/blog/pengertian-embedding/> (accessed Jul. 25, 2023).
- Nafan, M. Z., Bimantara, A. A., Larasati, A., Risondang, E. M., & Nugraha, N. A. S. (2019). Sentiment analysis of cyberbullying on instagram user comments. *Journal of Data Science and Its Applications*, 2(1), 38-48.
- Napoleoncat. (2022). Instagram Users in Indonesia. *napoleoncat.com*. (Online). <https://napoleoncat.com/stats/instagram-users-in-indonesia/2022/12/> (accessed Jun. 20, 2023).
- Olah, C. (2015). Understanding lstm networks. *colah.github.io*. (Online). <http://colah.github.io/posts/2015-08-Understanding-LSTMs/> (accessed Jul. 23, 2023).
- Pahlevi, R. (2022). “Jumlah Anak Korban Perundungan di Media Sosial (2016-2020)”. *databoks.katadata.co.id*. (Online). <https://databoks.katadata.co.id/datapublish/2022/07/29/ratusan-anak-jadi-korban-bullying-di-media-sosial-sejak-2016> (accessed Mar. 11, 2023).
- Pane, M. D. C. (2023). Body Shaming, Kenali Tanda-Tanda Perilakunya. *alodokter.com*. (Online). <https://www.alodokter.com/body-shaming-apakah-kamu-melakukannya> (accessed May. 11, 2024).
- Pardede, M. C. N. (2022). *Identifikasi Komentar Body Shaming pada Media Sosial Menggunakan Term Frequency Chi-Square dan Random Forest Classifier* (Doctoral dissertation, Universitas Sumatera Utara).

- Paren, S. 2022. “Membaca Statistic Tentang Kasus Bullying di Indonesia”. *depoedu.com*. (Online). <https://www.depoedu.com/2022/12/13/edu-talk/membaca-statistik-tentang-kasus-bullying-di-indonesia/> (accessed Mar. 10, 2023).
- Pasaribu, R. I. (2022). *Identifikasi Hate Speech pada Media Sosial Twitter Menggunakan Backpropagation Artificial Neural Network* (Doctoral dissertation, Universitas Sumatera Utara).
- Pratama, K. V. (2022). “Instagram, Media Sosial Pemicu Cyberbullying Tertinggi”. *tekno.kompas.com*. (Online). <https://tekno.kompas.com/read/2021/03/29/07164137/instagram-media-sosial-pemicu-cyberbullying-tertinggi> (accessed Mar. 11, 2023).
- Prijono, B. (2018). Pengenalan Long Short Term Memory (LSTM) dan Gated Recurrent Unit (GRU) – RNN Bagian 2. *indoml.com*. (Online). <https://indoml.com/2018/04/13/pengenalan-long-short-term-memory-lstm-dan-gated-recurrent-unit-gru-rnn-bagian-2/> (accessed July. 25, 2023).
- Purnamasari, N. M. G. D., Fauzi, M. A., Indriati, I., & Dewi, L. S. (2018). Identifikasi Tweet Cyberbullying pada Aplikasi Twitter menggunakan Metode Support Vector Machine (SVM) dan Information Gain (IG) sebagai Seleksi Fitur. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(11), 5236-5332.
- Raudhoti, L. A. H., Herdiani, A., & Romadhony, A. (2020). Identifikasi Cyberbullying pada Kolom Komentar Instagram dengan Metode Support Vector Machine dan Semantic Similarity. *Journal of Computer Science and Informatics Engineering (J-Cosine)*, 4(1), 1-8.
- Rumaisa, F., Puspitarani, Y., Rosita, A., Zakiah, A., & Violina, S. (2021). Penerapan Natural Language Processing (NLP) di bidang pendidikan. *Jurnal Inovasi Masyarakat*, 1(3), 232-235.
- Sabharwal, N., Agrawal, A., Sabharwal, N., & Agrawal, A. (2021). Introduction to word embeddings. *Hands-on Question Answering Systems with BERT: Applications in Neural Networks and Natural Language Processing*, 41-63.
- Soyusiawaty, D., & Haspiyan, R. (2015, July). Aplikasi Kamus Bahasa Indonesia–Bahasa Sasak Berbasis WAP. In *Seminar Nasional Informatika (SEMNASIF)* (Vol. 1, No. 5).
- Swim, J. K., & Hyers, L. L. (2009). 20 Sexism. *Handbook of Prejudice, Stereotyping, and Discrimination*, 407.

- Syarif, R. D., Herdiani, A., & Astuti, W. (2019). Identifikasi Cyberbullying Pada Komentar Instagram Menggunakan Metode Lexicon-based Dan Naïve Bayes Classifier (studi Kasus: Pemilihan Presiden Indonesia Tahun 2019). *eProceedings of Engineering*, 6(2).
- Thaliakr. (2016). Everyday Misogyny: 122 Subtly Sexist Words About Women (And What To Do About Them). *sacraparental.com*. (Online).
<https://sacraparental.com/2016/05/14/everyday-misogyny-122-subtly-sexist-words-women/> (accessed May. 25, 2024).
- Thompson, N. (1995). Men and anti-sexism. *The British Journal of Social Work*, 25(4), 459-475.
- Trihapsari, E. (2016). Klasifikasi Cyber Bullying Pada Media Sosial Twitter Dengan Menggunakan Algoritma Naïve Bayes. Tesis. Institut Teknologi Sepuluh Nopember.
- Weiss, S. M., Indurkha, N., Zhang, T., & Damerau, F. (2010). *Text mining: predictive methods for analyzing unstructured information*. Springer Science & Business Media.
- Willard, N. E. (2007). *Cyberbullying and cyberthreats: Responding to the challenge of online social aggression, threats, and distress*. Research press. vol. 14, no. 5, pp.1-13.