JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 8, No. 2, Juni 2023, Pp. 551-560

PENGIMPLMENTASIAN ALGORITMA LONG SHORT-TERM MEMORY UNTUK MENDETEKSI UJARAN KEBENCIAN PADA APLIKASI TWITTER

Renaldo Yosia Rafael*[1)], Fransiskus Adikara[2)]
1. Universitas Bunda Mulia, Indonesia
2. Universitas Bunda Mulia, Indonesia

Article Info
Kata Kunci: Long short-term memory, LSTM, Machine learning, Text pre-processing, Twitter, Ujaran kebencian

Keywords: Long short-term memory, LSTM, Machine learning, Text pre-processing, Twitter, Hate Speech

Article history:
Received 18 January 2023
Revised 25 January 2023
Accepted 20 February 2023
Available online 1 June 2023

DOI :
https://doi.org/10.29100/jipi.v8i2.3490

* Corresponding author.
Corresponding Author
E-mail address:
s32180117@student.ubm.ac.id

ABSTRAK
Memasuki tahun 2022, jumlah pengguna internet di Indonesia sudah mencapai angka 204.7 juta pengguna, dimana se-bagian besar penggunaan internet adalah untuk media social. Seiring dengan ting-ginya pengguna media sosial, Direktorat Tindak Pidana Siber Bareskrim Polri mendapati 89 konten media sosial terverifikasi mengandung ujaran kebencian selama periode Februari-Maret 2021, dimana konten terban-yak berasal dari aplikasi Twitter. Oleh karena itu dilakukanlah penelitian dengan mengimplementasikan Machine learning untuk mendeteksi uja-ran kebencian pada ap-likasi Twitter menggunakan metode Long short-term memory. Penyerapan data Twitter dilakukan dengan mengimple-men-tasikan Tweepy Library oleh Muhammad Okky Ibrohim yang di-akses melalui Kaggle selama sekitar 7 bulan, mulai 20 Maret 2018 hingga 10 September 2018. Tujuan penyerapan data dengan waktu yang lama adalah untuk mendapatkan lebih banyak pola tulisan twit. Data yang telah melalui text processing kemudian dibuat menjadi token-token yang meru-pakan rangkaian nilai integer. Kemudian model LSTM dibangun dengan mengkompilasikan input layer, LSTM layer, dan out-put layer, untuk nantinya dilatih dengan data training yang telah dipisahkan dari dataset. Peneliti menemukan hasil dari pelatihan model menunjukan accuracy sebesar 95.74% dan nilai loss 0.3463. Saat model yang telah dilatih digunakan untuk membuat prediksi terhadap data test, peneliti mendapatkan nilai accuracy sebesar 90% yang mengindi-kasikan model telah melakukan prediksi secaara akurat. Berdasarkan per-forma model dalam mendeteksi ujaran keben-cian, peneliti dapat me-nyimpulkan bahwa deteksi ujaran kebencian pada Twitter dapat dil-akukan dengan penggunaan Machine learning dan algoritma Long short-term memory(LSTM) dengan tingkat akurasi yang cukup tinggi.

ABSTRACT
Entering 2022, the number of internet users in Indonesia has reached 204.7 million users, where most of the internet us-age is for social media. Along with the high number of social media users, the Directorate of Cyber Crime of the Criminal Investigation Unit of the National Police found 89 verified social media content containing hate speech during the Febru-ary-March 2021 period, where the most content came from the Twitter application. Therefore, a research was conducted by implement-ing Machine learning to detect hate speech on the Twitter application using the Long short-term memory meth-od. Twitter data absorption was carried out by implementing the Tweepy Library by Muhammad Okky Ibrohim which was accessed via Kaggle for about 7 months, from March 20, 2018 to September 10, 2018. Data that has gone through text pro-cessing is then made into tokens which are a series of integer values. Then the LSTM model is built by compiling the input layer, LSTM layer, and output layer, to be trained later with training data that has been sep-arated from the dataset. The researcher found that the results of the model training showed an accuracy of 95.74% and a loss value of 0.3463. When the trained model is used to make predictions on the test data, the re-searcher gets an accuracy value of 90% which indicates the model has made accurate predictions. Based on the model's performance in detect-ing hate speech, research-ers can conclude that hate speech detection on

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 8, No. 2, Juni 2023, Pp. 551-560

Twitter can be done using Machine learning and Long short-term memory (LSTM) algorithms with a fairly high level of accuracy.

I. PENDAHULUAN

Pada tahun 2022, jumlah pengguna internet di Indonesia sudah mencapai angka 204.7 juta pengguna, atau sekitar 73,7% dari total jumlah penduduk Indonesia. Dari angka tersebut, sebagian besar dari penggunaan internet ditujukan untuk berinteraksi di media sosial. Hal ini dapat dilihat dari tingginya angka pengguna media sosial di Indonesia yaitu mencapai 191.4 juta pengguna, atau setara dengan 68.9% dari total jumlah penduduk Indonesia. Adapun platform media sosial yang dinyatakan paling favorit oleh pengguna dengan rentang umur 16-64 tahun adalah *Whatsapp*, *Instagram*, *Facebook* dan *Twitter* [1].

Media sosial merupakan sarana bagi pengguna internet atau yang biasa disebut dengan sebutan warganet untuk dapat berinteraksi secara daring. Pada aplikasi *Twitter*, warganet dapat berkomunikasi melalui twit yang diunggah. Twit yang diunggah dapat bersifat positif atau negatif dinilai dari konten yang terkandung dalam twit tersebut. Twit yang berisi komentar yang negatif menjadi masalah karena biasanya mengandung unsur SARA dan/atau ujaran kebencian, dimana tindakan serupa dapat berakibat sanksi hukum bagi penulisnya [2].

Direktorat Tindak Pidana Siber(Dit. Tipisiber) Bareskrim Polri mendapati 89 konten media sosial terverifikasi mengandung ujaran kebencian selama periode 23 Februari - 11 Maret 2021, konten terbanyak berasal dari *Twitter*. Berdasarkan data dari Dit. Tipisiber Bareskrim Polri, pada periode itu ada 125 konten yang diajukan untuk diberikan peringatan. Konten yang diberi peringatan didominasi oleh platform *Twitter* yaitu 79 konten, *Facebook* 32 konten, *Instagram* 8 konten, YouTube 5 konten dan *Whatsapp* satu konten [3].

Ujaran kebencian merupakan fenomena dalam bidang Bahasa yang bertolak belakang dan melanggar konsep kesantunan berbahasa dan etika berkomunikasi. Sifat '*openness of media*' atau keterbukaan informasi di media sosial inilah salah satu yang memicu tingginya kecenderungan masyarakat untuk melakukan ujaran kebencian. Ketersediaan fasilitas komentar untuk pembaca pada media yang berbasis elektronik membuat pengguna memiliki kesempatan. Hal itu menyebabkan hubungan antara peneliti dan pembaca menjadi resiprokal, bisa, dan mudah untuk saling mengomentari [4].

Rasisme sekarang ini menjadi permasalahan dunia yang masih belum bisa teratasi sepenuhnya. Penjelasan mengenai isu rasisme tidak hanya dijelaskan melalui media konvensional, saat ini isu rasisme telah masuk kedalam media elektronik seperti film dan media social [5].

Seksisme adalah tindakan atau kata-kata yang mengarah kepada kebencian atau diskriminasi berdasarkan pada jenis kelamin seseorang. Sikap sexist mungkin berakar dari stereotip tradisional atau peran jenis kelamin (gender roles). Kontroversi Gender mengungkapkan beberapa perbedaan substansial dalam kemampuan fisik, keterampilan membaca dan menulis, agresi, dan pengaturan diri, hanya ada sedikit perbedaan dalam kemampuan matematika, dan ilmu pengetahuan [6].

Machine learning merupakan cabang atau aplikasi dari kecerdasan buatan yang berfokus untuk membuat sistem yang terus belajar dari data dan meningkatkan akurasinya dari waktu ke waktu. Salah satu sub-bagian dari *Machine learning* yaitu *Text analytic* memiliki algoritma-algoritma yang dapat melakukan pengenalan atau pengelompokan terhadap suatu objek teks. *Text analytic* ini dapat dimanfaatkan dalam mengatasi kasus ujaran kebencian dalam media sosial melalui kemampuannya dalam mendeteksi cyber bullying, bahasa kasar, maupun cyber hate [7].

Long Short Term Memory adalah bentuk modifikasi dari algoritma *Recurrent Neural Network*(RNN), yang mampu mengatasi kelemahan RNN yaitu mempelajari dependensi jangka panjang. LSTM bekerja dengan sangat baik pada berbagai macam masalah yang bersifat sequence, dan sekarang banyak digunakan. Metode LSTM adalah memberikan prediksi dengan menggunakan langkah demi langkah dari urutan(sequence) data. Keunggulan LSTM adalah cocok untuk prediksi time series dan ketahanannya dalam menangani data yang besar dan non-linier [8].

Berdasarkan masalah yang ada, dengan menggunakan algoritma *Long short-term memory*, peneliti akan mengembangkan sistem berbasis *Machine learning* yang dapat mendeteksi ujaran kebencian yang dibatasi pada kategori rasisme dan seksisme pada twit berbahasa Indonesia. Selain itu penelitian ini juga bertujuan untuk menguji tingkat akurasi dari model *Long short-term memory* dalam mendeteksi ujaran kebencian.

Bahasa pemrograman yang dipilih adalah Python, karena memiliki beberapa keunggulan seperti *readability*, *efficiency*, *extensible* & *embeddable*, dan memiliki dukungan komunitas. *Readibility* yang dimaksud adalah Python memiliki source code yang sederhana, command yang merupakan sebuah kata dalam Bahasa inggris, sehingga mudah ditulis, mudah diingat dan juga digunakan ulang. *Efficiency* yang dimaksud adalah, Python memiliki library yang lengkap dan code yang lebih sederhana bila dibandingkan dengan kode yang ditulis dengan bahasa pemrograman lainnya seperti misalnya Java, C, C# maupun C++. *Extensible* & *embeddable*, untuk

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 8, No. 2, Juni 2023, Pp. 551-560

mengembangkan aplikasi dengan performa tinggi. Program yang dikembangkan Python juga dapat dioperasikan pada hampir semua sistem operasi baik Windows, Linux, Mac OS, Unix, dan juga sistem operasi pada perangkat lunak berbasis mobile seperti Android atau IOS [9].

II. METODE PENELITIAN

Secara garis besar, metode yang digunakan dalam penelitian ini, antara lain sebagai berikut :

A. Metode Pengumpulan

Kerangka kerja yang dibuat dalam penelitian ini membutuhkan bermacam-macam informasi untuk menjadi langkah perencanaan penelitian. Untuk mendapatkan data awal penelitian, maka dilakukan langkah berikut untuk mendapatkan data-datanya yaitu:

1) Studi literatur, dengan membaca buku referensi dan tulisan yang berhubungan dengan materi *Machine learning*.

2) Streaming data pada *Twitter*, Dalam penelitian ini, data yang digunakan adalah data twit yang didapatkan dari hasil penarikan pada server *Twitter* menggunakan *Twitter Search* API dengan mengimplementasikan Tweepy Library oleh Muhammad Okky Ibrohim yang diakses melalui Kaggle mulai 20 Maret 2018 hingga 10 September 2018 (contoh data pada gambar 1).


preprocessed_indonesian_toxic_tweet.csv (1.86 MB)

Tweet	HS	Abusive	HS_Individual	HS_
di saat semua cowok berusaha melacak perhatian gue kamu lantas remehkan perhatian yang gue kasih kh...	1	1	1	0
siapa yang telat memberi tau kamu edan sarap gue bergaul dengan cigax jifla colis sama siapa itu li...	0	1	0	0
41 kadang aku berpikir kenapa aku tetap percaya pada tuhan padahal aku selalu jatuh berkali kali kad...	0	0	0	0

Gambar 1. Sample Data yang Digunakan

B. Analisis Permasalahan

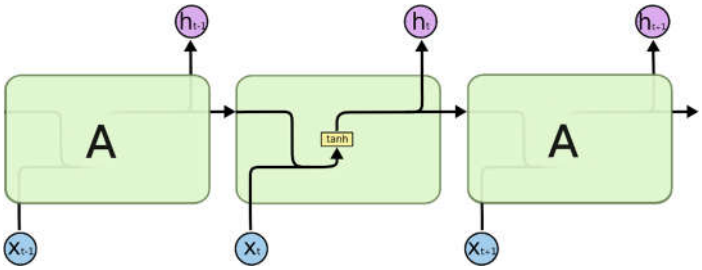
Setelah peneliti mengumpulkan data dari studi pustaka dan dataset, maka peneliti mendapatkan masalah yaitu bagaimana untuk melakukan deteksi ujaran kebencian pada aplikasi *Twitter* dengan model *Machine learning*, sehingga dapat diketahui klasifikasi sebuah twit termasuk ujaran kebencian atau tidak.

Algoritma yang digunakan peneliti untuk mendeteksi ujaran kebencian pada *Twitter* berbahasa Indonesia adalah algoritma *Long short-term memory*. Pemilihan *Long short-term memory* untuk kasus deteksi ujaran kebencian ini didasari oleh kemampuan *Long short-term memory* untuk mempelajari data secara sekuensial, sehingga cocok untuk kasus deteksi ujaran kebencian, dimana data yang diolah merupakan data yang berbentuk kata-kata dan bersifat sekuensial.

Algoritma Long short-term memory Jaringan Long Short Term Memory adalah bentuk modifikasi dari algoritma Recurrent Neural Network(RNN), yang mampu mengatasi kelemahan RNN yaitu mempelajari dependensi jangka panjang. LSTM bekerja dengan sangat baik pada berbagai macam masalah yang bersifat sequence, dan sekarang banyak digunakan. Metode LSTM adalah memberikan prediksi dengan menggunakan langkah demi langkah dari urutan(sequence) data. Keunggulan LSTM adalah cocok untuk prediksi time series dan ketahanannya dalam menangani data yang besar dan non-linier.[12]

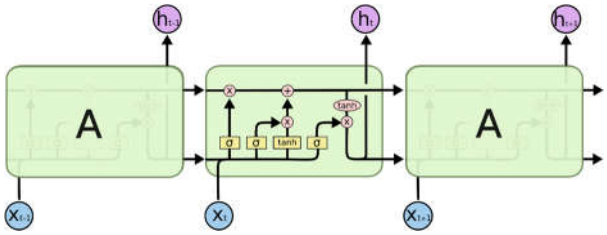
Perbedaan dari LSTM dan RNN terletak pada lapisan rantainya. Jika RNN memungkinkan satu neuron untuk memproses satu input data pada satu output data, LSTM tidak berlaku demikian. LSTM memiliki berbagai gerbang yang dapat menambah kumpulan informasi dan kemudian menggabungkannya.

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 8, No. 2, Juni 2023, Pp. 551-560



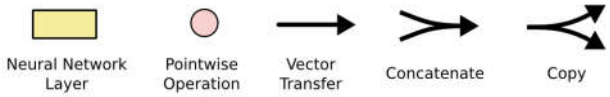
Gambar 2 Rantai Modul RNN

Ada empat gerbang dalam sistem LSTM, yakni forget gate, input gate, input modulation gate, serta output gate. Keempat gerbang tersebut mempunyai fungsi dan tugasnya masing-masing dalam mengumpulkan, mengklasifikasi, dan memproses data. Selain empat gerbang tersebut, LSTM juga memiliki internal cell state yang berfungsi untuk menyimpan informasi pilihan dari unit sebelumnya. Input gate pada modul LSTM berperean untuk mengendalikan sejauh mana nilai baru akan berjalan ke dalam cell, forget gate mengendalikan apakah nilai akan tetap berada dalam cell, dan output gate berperan untuk mengendalikan nilai dalam cell yang digunakan untuk menghitung keluaran dari unit LSTM.



Gambar 3 Rantai modul LSTM

Notasi :



Setiap garis membawa seluruh vektor, dari output satu node ke input lainnya. Lingkaran merah muda mewakili operasi titik, seperti penambahan vektor, sedangkan kotak kuning adalah lapisan jaringan saraf yang dipelajari. Penggabungan garis menunjukkan penggabungan, sedangkan percabangan garis menunjukkan kontennya disalin dan salinannya pergi ke lokasi yang berbeda.

Data yang masuk pada forget gate akan diolah sesuai informasinya dan kemudian dilakukan pemilihan data yang mana yang akan disimpan pada memory cell. Fungsi aktivasinya menggunakan sigmoid. Persamaan (1) menggambarkan prinsip kerjanya, sedangkan input gates memiliki 2 gate yang menggunakan fungsi aktivasi sigmoid untuk memperbarui informasi dan menggunakan fungsi aktivasi tanh yang akan menyimpan nilai baru di memory cell. Hal ini dapat digambarkan pada persamaan (2) dan (3).

$$f_t = \sigma\left(W_f.\left[h_{t-1}, x_t\right] + b_t\right) \quad (1)$$

$$i_t = \sigma\left(W_i.\left[h_{t-1}, x_t\right] + b_i\right) \quad (2)$$

$$\hat{c}_t = \tanh\left(W_c.\left[h_{t-1}, x_t\right] + b_c\right) \quad (3)$$

Persamaan (4) adalah hasil gabungan nilai pada input gate. Forget gate akan menggantikan nilai memory cell oleh cell gates. Pada output gates juga terdapat 2 gate yaitu untuk memutuskan nilai yang akan dikeluarkan dengan fungsi aktivasi sigmoid dan menyimpan nilai dengan memakai fungsi aktivasi tanh. Hal ini dirumuskan pada persamaan (5) dan (6)[12].

$$c_t = f_t * c_{t-1} + i_t * \hat{c}_t \quad (4)$$

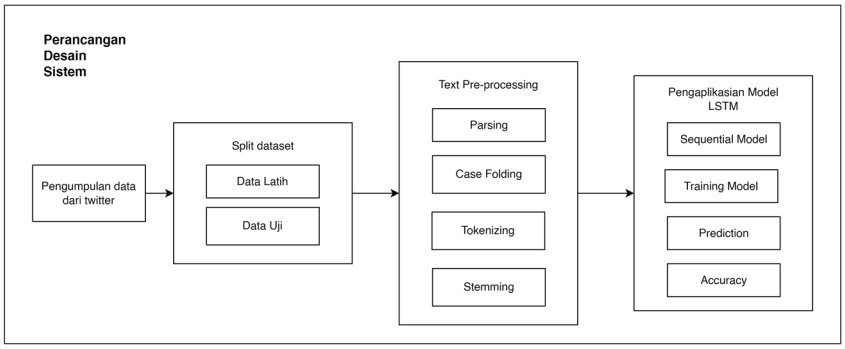
$$o_t = \sigma\left(W_o.\left[h_{t-1}, x_t\right] + b_o\right) \quad (5)$$

$$h_t = o_t \tanh\left(c_t\right) \quad (6)$$

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 8, No. 2, Juni 2023, Pp. 551-560

C. *Desain dan Implementasi*

Peneliti melakukan perancangan pada sistem pendeteksi ujaran kebencian pada twit berbahasa Indonesia, dimana tahap-tahapnya secara garis besar dapat dilihat dalam bentuk diagram alur pada Gambar 1.



Gambar 2. Diagram Alur Perancangan Sistem

D. *Pengujian dan Akurasi*

Pengujian yang dilakukan dalam penelitian ini bertujuan untuk mengetahui tingkat akurasi proses deteksi ujaran kebencian dari model yang telah dibangun. Pengujian pada penelitian ini dilakukan dengan membuat confusion matrix. Melalui tabel confusion matrix dapat diketahui nilai dari *Accuracy*, *Precision*, *Recall* dan *F1-Score*. *Confusion matrix* adalah sebuah metode pengujian untuk mengukur performa sebuah model pada *Machine learning*. Pada dasarnya confusion matrix memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh sistem (model) dengan hasil klasifikasi sebenarnya. Confusion Matrix berbentuk tabel matriks 2x2 yang berisi empat istilah sebagai representasi nilai prediksi dan nilai aktual yang berbeda. Keempat istilah tersebut adalah *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False Negative* (FN) [10].

Seperti yang telah dijelaskan bahwa confusion matrix dapat mengukur performa(performance metrics) dari sebuah model. Perfomance metrics yang akan diukur pada penelitan ini adalah sebagai berikut[13]:

a. Accuracy

Accuracy (Akurasi) menggambarkan seberapa akurat model dapat melakukan klasifikasi dengan benar. Maka, dapat disimpulkan bahwa akurasi merupakan rasio prediksi benar (positif dan negatif) disbanding keseluruhan data. Dengan kata lain, akurasi merupakan nilai yang merepresentasikan tingkat kedekatan antara nilai dari prediksi dengan nilai yang aktual. Rumus perhitungan akurasi dapat dilihat pada persamaan (1).

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

b. Precision

Precision (Presisi) menggambarkan tingkat keakuratan antara data yang diminta dengan hasil prediksi yang diberikan oleh model. Maka, presisi merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Dari semua kelas positif yang telah diprediksi bernilai benar, berapa banyak data yang benar-benar positif. Untuk menghitung presisi dapat dilihat pada persamaan (2)

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

c. *Recall*

Recall menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi. Maka, *recall* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Nilai recall dapat diperoleh dengan persamaan (3)

$$\text{recall} = \frac{TP}{TP+FN} \quad (3)$$

III. HASIL DAN PEMBAHASAN

Untuk melakukan deteksi ujaran kebencian dengan model LSTM dilakukan langkah – langkah yaitu Langkah 1, mengolah dataset. Untuk menyesuaikan dataset dengan kebutuhan penelitian ini, maka peneliti melakukan pembersihan dataset dengan menggunakan *Function* drop() pada Python untuk menghilangkan kolom dan baris data yang tidak berhubungan dengan rasisme dan seksisme. Dataset pada mulanya memiliki 13 kolom seperti pada Tabel 1, setelah melalui pembersihan, kolom yang tersisa hanya 4 kolom seperti pada Tabel 2.

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 8, No. 2, Juni 2023, Pp. 551-560

TABEL I
POTONGAN DATASET SEBELUM DIOLAH

No.	Tweet	HS	Abu-sive	HS_ Individ-ual	HS_ Group	HS_ Reli-gion	HS_ Race	HS_ Physi-cal	HS_ Gen-der	HS_ Other	HS_ Weak	HS_ Moder-ate	HS_ Strong
1	- disaat semua cowok berusaha melacak perhatian gue. loe lantas remehkan perhatian yg gue kasih khusus ke elo. basic elo cowok bego												
2	RT USER: USER siapa yang telat ngasih tau elu?edan sarap gue bergaul dengan cigax jifla calis sama siapa noh licew juga												
3	41. Kadang aku berfikir, kenapa aku tetap percaya pada Tuhan padahal aku selalu jatuh berkali-kali. Kadang aku merasa Tuhan itu ninggalkan aku sendirian. Ketika orangtuaku berencana berpisah, ketika kakakku lebih memilih jadi Kristen.												
4	USER USER AKU ITU AKU\n\nKU TAU MATAMU SIPIT TAPI DILIAT DARI MANA ITU AKU												
5	USER USER Kaum cebong kapir udah keliatan dongoknya dari awal tambah dongok lagi hahahah												

TABEL II
POTONGAN DATASET SESUDAH DIOLAH

No.	Tweet	HS	HS_Race	HS_Gender
1	USER: USER siapa yang telat ngasih tau elu?edan sarap?? gue bergaul dengan cigax jifla calis sama siapa noh licew juga	0	0	0
2	41. Kadang aku berfikir, kenapa aku tetap percaya pada Tuhan padahal aku selalu jatuh berkali-kali. Kadang aku merasa Tuhan itu ninggalkan aku sendirian. Ketika orangtuaku berencana berpisah, ketika kakakku lebih memilih jadi Kristen.	0	0	0
3	USER USER AKU ITU AKU\n\nKU TAU MATAMU SIPIT TAPI DILIAT DARI MANA ITU AKU	0	0	0
4	deklarasi pilkada 2018 aman dan anti hoax warga dukuh sari jabon	0	0	0
5	Gue baru aja kelar re-watch Aldnoah Zero!!! paling kampret emang endingnya!?? 2 karakter utama cowonya kena friendzone bray! XD	0	0	0

Langkah 2, melakukan *text pre-processing* yang terdiri dari beberapa proses yaitu penghapusan simbol dan emoji dengan mendefinisikan sebuah class berisi kompilasi list Unicode dari emoji dan simbol yang ingin dihilangkan kemudian menghapus class tersebut.

Selain menghapus simbol dan emoji, juga perlu menghapus tanda baca(punctuation) dengan mendefinisikan memanfaatkan *Function* str.maketrans(). Setelah itu dilakukan proses *Tokenizing*, dimana terjadi pemecahan dari satu kalimat menjadi kata per kata. Dalam penelitian ini, *Tokenizing* dilakukan dengan membuat text array yang berisikan masing-masing kata dari tiap kalimat. Kemudian diberlakukan kondisi dimana selama value bukan

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 8, No. 2, Juni 2023, Pp. 551-560

berupada digit, dan memiliki length lebih dari 3 maka text akan masuk ke dalam text array untuk menghilangkan angka dan kata sambung.

Tweet sebelum melalui proses *text processing* dapat dilihat pada Tabel 3, Tweet sesudah melalui proses *text processing* dapat dilihat pada Tabel 4.

TABEL III
TWIT SEBELUM *TEXT PROCESSING*

No.	Twit
1	USER USER AKU ITU AKU\n\nKU TAU MATAMU SIPIT TAPI DILIAT DARI MANA ITU AKU'
2	Gue baru aja kelar re-watch Aldnoah Zero!!! paling kampret emang endingnya! 2 karakter utama cowonya kena friendzone bray! XD URL
3	Rizieq shihab fpi jancok asu kontol tempek anjing babi bajingan bangsat lonte balon banci bencong taek cabul pengecut cok teroris bubarkan ormas fpi #HTIMakar'
4	Ruhut Sitompul: Prabowo Jangan Omdo URL
5	USER Pak Recep.....anda salah, itu gubernur pakkkk....bukan presiden ., presiden kami lagi di Got.... ; Nih liat kalo gak percaya...

TABEL IV
TWIT SESUDAH *TEXT PROCESSING*

No.	Twit
1	user user akunnku matamu sipit tapi diliat dari mana
2	baru kelar rewatch aldnoah zero paling kampret emang endingnya karakter utama cowonya kena friendzone bray
3	rizieq shihab jancok kontol tempek anjing babi bajingan bangsat lonte balon banci bencong taek cabul pengecut teroris bubarkan ormas htimakar
4	ruhut sitompul prabowo jangan omdo
5	user recepanda salah gubernur pakkkkbukan presiden presiden kami lagi liat kalo percaya

Langkah 3, membagi dataset menjadi *data training* dan *data test* menggunakan *Function* train_test_split() dengan proporsi *data training* dan *data test* 5:1. *Data training* akan digunakan untuk melatih model LSTM, sedangkan *data test* akan digunakan untuk menguji performa LSTM. Kemudian dari *data training* yang ada dibagi lagi menjadi *data training* dan data validation dengan proporsi *data training* dan data validation adalah 5:1 juga.

Langkah 4, mempersiapkan *sequential model* yang dimulai dengan proses *Tokenizing* yaitu transformasi kata-kata pada *data train* menjadi rangkaian integer yang akan merepresentasikan tiap kata di dalam sebuah twit. Token-token yang dibentuk kemudian digabungkan ke dalam numpy array yang dibuat untuk masing-masing twit.

Karena jumlah kata dalam twit pasti berbeda-beda, maka dilakukan *Padding*, untuk menyamakan length array dari integer-integer yang merupakan hasil *Tokenizing* sebelumnya. Hasil array setelah *Padding* dapat dilihat pada Gambar 2. Nilai 0 pada array merupakan hasil *Padding* yang berfungsi menyamakan jumlah kata untuk setiap twitnya. Batas maksimal jumlah kata dalam sebuah twit ditetapkan sebanyak 50 kata.

```
=======Train Dataset=======
tf.Tensor(
[   2    2   48 4213  318  827 9347    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0], shape=(50,), dtype=int32) tf.
=======Validation Dataset=======
tf.Tensor(
[ 118  623  165  588 1108  464  764 1186 1109  798   24  728  173  458
   13   35  487  990  991   86    1  729  698    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0], shape=(50,), dtype=int32) tf.
=======Test Dataset=======
tf.Tensor(
[ 118  623  165  588 1108  464  764 1186 1109  798   24  728  173  458
   13   35  487  990  991   86    1  729  698    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0], shape=(50,), dtype=int32) tf.
```

Gambar 2. Array Setelah Proses Padding

Langkah 5 adalah membangun model *Long short-term memory* yang dimulai dari menentukan fitur maksimal untuk dipelajari, menentukan embedding dimension dan Panjang sequence seperti perintah pada Gambar 3.

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 8, No. 2, Juni 2023, Pp. 551-560

```
#model preparation
max_features =50000
embedding_dim =12
sequence_length = maxlen
```

Gambar 3. Persiapan sebelum membuat model LSTM

Kemudian selanjutnya peneliti melakukan definisi *sequential model* dan kemudian menambahkan berbagai layers ke dalamnya. Layer pertama adalah embedding layer(tf.keras.layers.Embedding). Penempatan sebuah kata dalam ruang vektor didasarkan pada kata-kata yang berada di sekitar kata tersebut saat digunakan, seperti yang diilustrasikan pada Gambar 4.

Gambar 4. Cara Kerja Word Embedding[11]

Lapisan berikutnya adalah lapisan LSTM yaitu lapisan dengan jumlah neuron setara dengan besar embedding_dim dan dropout yang dibuat sebesar 0.2 seperti yang dapat dilihat pada Gambar 5.

```
model.add(tf.keras.layers.LSTM
          (embedding_dim, dropout=0.2,
           recurrent_dropout=0.2,
           return_sequences=True, \
           kernel_regularizer = regularizers.l2(0.005), \
           bias_regularizer = regularizers.l2(0.005)))
```

Gambar 5. Perintah Untuk Membangun Layer LSTM

Dropout bekerja secara acak mengatur bagian pinggir dari hidden unit untuk kembali ke 0 pada setiap pembaharuan fase training agar terhindar dari bias. return_sequences=True adalah parameter penting saat menggunakan beberapa layer LSTM karena memungkinkan output dari layer LSTM sebelumnya digunakan sebagai input ke layer LSTM berikutnya. Jika tidak ditetapkan nilai true, maka layer LSTM berikutnya tidak akan mendapatkan input.

```
model.add(tf.keras.layers.Dense
          (8, activation='relu', \
           kernel_regularizer = regularizers.l2(0.001),\
           bias_regularizer = regularizers.l2(0.001),))
model.add(tf.keras.layers.Dropout(0.4))

model.add(tf.keras.layers.Dense
          (1, activation='sigmoid'))
```

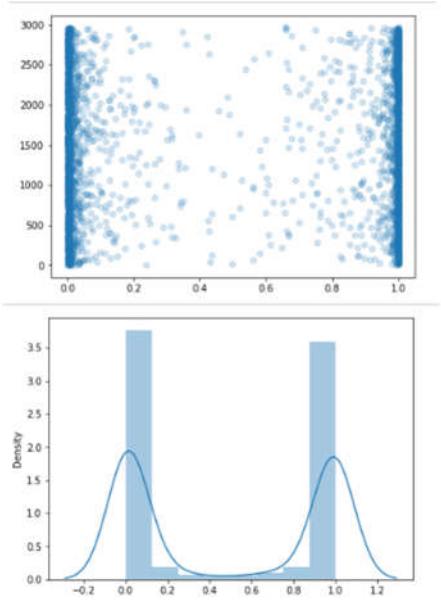
Gambar 6. Layer Output Pada Model LSTM

Pada Gambar 8, Layer Dense yang terakhir adalah output layer yang memiliki 1 sel output. Fungsi aktivasi ‘softmax’ digunakan di layer terakhir yang berguna untuk mengubah nilai vektor n menjadi nilai vektor n yang berjumlah 1. Nilai input bisa positif, negatif, nol, atau lebih besar dari satu, tetapi softmax mengubahnya menjadi nilai antara 0 dan 1, sehingga dapat diinterpretasikan sebagai probabilitas. Kemudian layer-layer yang ada dikompilasikan dan membentuk model LSTM.

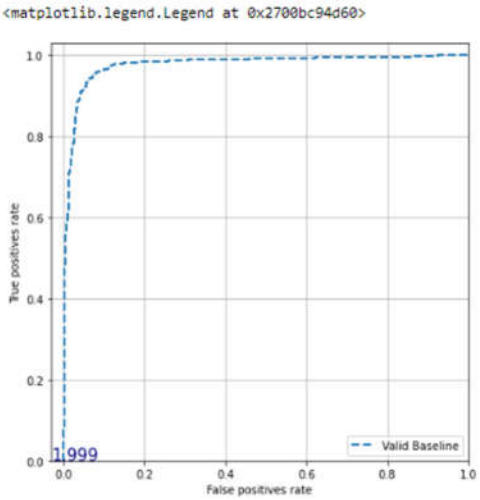
Model LSTM yang sudah dibangun kemudian dilatih dengan *data train* dengan 15 epoch dan 1024 batch. Setelah mencapai epoch terakhir didapatkan accuracy 0.9574 dan nilai loss yang terus menurun hingga 0.3463.

Gambar 7 meunjukkan visualisasi data hasil predictions model LSTM terhadap *data test*. Dimana distribusi hasilnya sebagian besar terbagi di titik 0 dan 1. Visualisasi selanjutnya adalah berupa plot ROC(Receiver Operative Characteristics) yang menunjukkan kinerja model klasifikasi. Kurva ini menampilkan dua parameter: True Positive Rate dan False Positive Rate. Pada Gambar 8, dapat dilihat hasil True positive dari waktu ke waktu bertambah dan mendekati nilai 1.0.

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Journal homepage: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi
ISSN: 2540-8984
Vol. 8, No. 2, Juni 2023, Pp. 551-560



Gambar 7. Visualisasi Hasil Deteksi Ujaran Kebencian Dalam *Data Test*



Gambar 7. Kurva ROC Yang Menampilkan Nilai True Positive dan False Positive

Selanjutnya model akan dievaluasi akurasinya, seberapa baik model melakukan deteksi data akan dihitung. *Function* classification_report digunakan untuk menampilkan skor akurasi dalam bentuk nilai *precision*, *recall*, *f1-score* dan *support*. Hasil accuracy dari model LSTM pada *data test* dapat dilihat pada Tabel V.

TABEL V
ACCURACY MODEL LSTM

LABEL	Precision	Recall	F1-Score	Support
0	0.87	0.96	0.91	1510
1	0.95	0.85	0.89	1458
Average	0.91	0.90	0.90	2968

IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa deteksi ujaran kebencian pada *Twitter* dapat dilakukan dengan penggunaan *Machine learning* dan algoritma *Long short-term memory*(LSTM). Model LSTM yang telah dibangun dan dilatih serta digunakan untuk melakukan predictions menunjukan hasil akurasi yang cukup tinggi.

Tingkat akurasi model LSTM yang telah dilatih untuk melakukan predictions memiliki akurasi sebesar 95.74% dan loss sebesar 0.3463 pada *data training*. Pada *data testing*, model LSTM yang telah dilatih memiliki akurasi sebesar 90%.

DAFTAR PUSTAKA

[1] "Digital-2022-Indonesia-February-2022-v01_compressed.pdf." .
[2] A. Perwira, J. Dwitama, and K. Kunci, "Deteksi Ujaran Kebencian Pada Twitter Bahasa Indonesia Menggunakan Machine Learning : Reviu Literatur," vol. 1, pp. 31–39, 2021.
[3] "Virtual Police temukan konten ujaran kebencian terbanyak di Twitter - ANTARA News." https://www.antaranews.com/berita/2039738/virtual-police-temukan-konten-ujaran-kebencian-terbanyak-di-twitter (accessed Apr. 02, 2022).
[4] D. J. Ningrum, S. Suryadi, and D. E. Chandra Wardhana, "Kajian Ujaran Kebencian Di Media Sosial," *J. Ilm. KORPUS*, vol. 2, no. 3, pp. 241–252, 2019, doi: 10.33369/jik.v2i3.6779.
[5] A. Rafly, Z. Abidin, and F. O. Lubis, "Analisis Semiotika Mengenai Representasi Rasisme Terhadap Orang Kulit Hitam Dalam Film Blackkklansman," *Semiotika*, vol. 14, no. 2, pp. 135–147, 2020, [Online]. Available: http://journal.ubm.ac.id/.
[6] F. N. Rosyidah and N. Nurwati, "Gender dan Stereotipe: Konstruksi Realitas dalam Media Sosial Instagram," *Share Soc. Work J.*, vol. 9, no. 1, p. 10, 2019, doi: 10.24198/share.v9i1.19691.
[7] F. Alzami, R. A. Megantara, and ..., "Sentiment Analysis Untuk Deteksi Ujaran Kebencian Pada Domain Politik," *Sci. ...*, vol. 5, no. Sens 5, pp. 213–218, 2020, [Online]. Available: http://conference.upgris.ac.id/index.php/sens/article/view/1606%0Ahttp://conference.upgris.ac.id/index.php/sens/article/download/1606/711.
[8] Y. Sudriani, I. Ridwansyah, and H. A Rustini, "Long short term memory (LSTM) recurrent neural network (RNN) for discharge level prediction and forecast in Cimandiri river, Indonesia," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 299, no. 1, 2019, doi: 10.1088/1755-1315/299/1/012037.
[9] E. Retnoningsih and R. Pramudita, "Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python," *Bina Insa. Ict J.*, vol. 7, no. 2, p. 156, 2020, doi: 10.51211/biict.v7i2.1422.
[10] V.A.R.Barao, R.C.Coata, J.A.Shibli, M.Bertolini, and J.G.S.Souza, "No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析Title," *Braz Dent J.*, vol. 33, no. 1, pp. 1–12, 2022.
[11] "Word Embeddings Explained. What is Word Embedding ? | by Ashwin Prasad | Analytics Vidhya | Medium." https://medium.com/analytics-vidhya/word-embeddings-explained-62c046f7c79e (accessed Aug. 30, 2022).