

DETEKSI DEPRESI PENGGUNA TWITTER INDONESIA MENGUNAKAN LSTM-RNN

Ivan Dwi Nugraha¹, Yufis Azhar²

^{1,2}Informatika, Teknik, Universitas Muhammadiyah Malang

email: ivandwinugraha45@webmail.umm.ac.id¹, yufis@umm.ac.id²

Abstrak

Perkembangan media sosial yang semakin pesat, menciptakan keberagaman microblogging sosial, mendorong orang untuk mengekspresikan perasaan dan pendapat, Setiap tweet pada twitter mewakili ekspresi emosional penggunanya, hal ini dapat dijadikan studi kasus dalam mendeteksi kasus depresi dan menilai emosional pengguna twitter. Deteksi dan pencegahan depresi sangat sulit untuk dideteksi dan telah menjadi topik penelitian yang sangat menarik untuk diteliti sejak dekade terakhir. Beberapa penelitian yang berkaitan dengan twitter untuk mendeteksi pengguna media sosial yang mengalami depresi. Salah satu penelitian deteksi depresi melalui twitter menyimpulkan bahwa adanya korelasi antara keadaan depresi pengguna twitter terhadap sentiment yang mereka tweet menggambarkan keadaan depresi pengguna tersebut. Tujuan penelitian ini penelitian kami adalah untuk mengembangkan dan mengoptimalkan penelitian sebelumnya menggunakan metode yang berbeda yakni LSTM-RNN, dan mendeteksi depresi pada tweet twitter indonesia. Dataset yang digunakan berjumlah 5.494 baris tweet, dimana data kelas normal berjumlah 2.747 baris tweet dan data depresi berjumlah 2.747 baris tweet setelah dilakukan balancing data, dataset sebelum digunakan data dilakukan proses preprocessing terlebih dahulu sebelum masuk ke proses pelatihan. Hasil dari penelitian dengan menggunakan metode LSTM-RNN memperoleh nilai presisi, recall, dan F1-score diperoleh masing-masing 86%, 86%, dan 86%, sedangkan akurasi adalah 86%. Sistem deteksi ujaran depresi diharapkan dapat membantu menganalisa depresi masyarakat di media sosial.

Kata kunci: NLP, Klasifikasi Teks, LSTM-RNN, Deteksi Depresi

Abstract

The rapid development of social media, creates a diversity of social microblogging, encourages people to express feelings and opinions. Every tweet on Twitter represents the emotional expression of its users, this can be used as a case study in detecting cases of depression and assessing the emotions of twitter users. Detection and prevention of depression is very difficult to detect and has been a very interesting research topic for research since the last decade. Several studies related to twitter to detect social media users who are depressed. One study of depression detection through Twitter concluded that there is a correlation between the depressed state of Twitter users and the sentiments they tweet describe the depressed state of the user. The purpose of this study is to develop and optimize previous research using a different method, namely the LSTM-RNN, and to detect depression in Indonesian twitter tweets. The dataset used is 5,494 tweet lines, where normal class data is 2,747 tweet lines and depression data is 2,747 tweet lines after data balancing, the dataset before being used is preprocessed before entering the training process. The results of the study using the LSTM-RNN method obtained precision, recall, and F1-score values obtained respectively 86%, 86%, and 86%, while the accuracy was 86%. Depression speech detection system is expected to help analyze people's depression on social media.

Keywords : NLP, Text Classification, LSTM-RNN, Depression Detection

Diterima Redaksi: 25-07-2022 | Selesai Revisi: 19-12-2022 | Diterbitkan Online: 27-12-2022

DOI: <https://doi.org/10.23887/janapati.v11i3.50674>

PENDAHULUAN

Pada saat ini perkembangan teknologi membuat media sosial berkembang semakin pesat, menciptakan keberagaman microblogging sosial, mendorong orang untuk

mengekspresikan perasaan dan pendapat mereka kapanpun[1]. Pada januari 2022 pengguna media sosial di Indonesia berjumlah 191,4 juta. Jumlah pengguna media sosial di Indonesia pada awal tahun 2022 setara dengan 68,9 persen dari total populasi di Indonesia,

jumlah pengguna media sosial di Indonesia meningkat sebesar 21 juta atau 12,5 persen dengan rentang waktu tahun 2021 sampai 2022[2]. Salah satu faktor popularitas media sosial yang semakin meningkat adalah kesempatan pengguna untuk menerima atau membuat dan berbagi informasi publik dengan kemudahan yang ditawarkan dan di manapun[3]. Platform media social juga banyak menawarkan format data, termasuk data tekstual, gambar, video, suara, dan geolokasi[4]. Sifat media sosial yang terpisah membuat orang lebih cenderung mengekspresikan diri mereka secara terbuka daripada secara langsung[5].

Perkembangan social media menciptakan banyak sekali platform media sosial yang disediakan oleh pengembang, salah satunya yaitu twitter. Twitter merupakan salah satu platform media social dimana pengguna menggunakannya, sebagai sarana berbagi informasi dengan sesama pengguna twitter secara realtime[6]. Pengguna twitter dapat berbagi kisah, opini dan aktivitasnya kepada public, dengan mengirimkan cuitan atau tweet[6]. Setiap tweet mewakili ekspresi emosional penggunanya, hal ini dapat dijadikan studi kasus dalam mendeteksi kasus depresi dan menilai emosional pengguna twitter[7][8]. Depresi mempengaruhi lebih dari empat persen populasi dunia. Depresi jangka panjang adalah kondisi kesehatan serius yang sangat mempengaruhi kemampuan seseorang untuk berfungsi dalam kehidupan sehari-hari[9]. Deteksi dan pencegahan depresi sangat sulit untuk dideteksi dan telah menjadi topik penelitian yang sangat menarik untuk diteliti sejak dekade terakhir. Beberapa penelitian yang berkaitan dengan twitter untuk mendeteksi pengguna media sosial yang mengalami depresi. Salah satu penelitian deteksi depresi melalui twitter menyimpulkan bahwa adanya korelasi antara keadaan depresi pengguna twitter terhadap sentiment yang mereka tweet menggambarkan keadaan depresi pengguna tersebut[7].

Penelitian klasifikasi emosi pada data Twitter Indonesia[10] membandingkan beberapa metode machine learning, seperti Logistic Regression, SVM, dan Random Forest kemudian menggabungkannya dengan beberapa metode ekstraksi fitur, seperti Bag-of-Words (BOW), Word2Vec (WV), dan Fast-Text (FT). Didapatkan hasil skor F1 sebesar 68,39%. Kemudian, mereka menambahkan tiga metode, yaitu Emotion Lexicons, Orthographic, dan POS Tag. Hasil yang didapatkan sedikit lebih tinggi dari pada pengujian sebelumnya, dengan skor

F1 sebesar 69,73%. Dari penelitian ini dihasilkan dataset yang disebut "Indonesian Twitter Emotion Dataset" sebagai hasil dari penelitian[10].

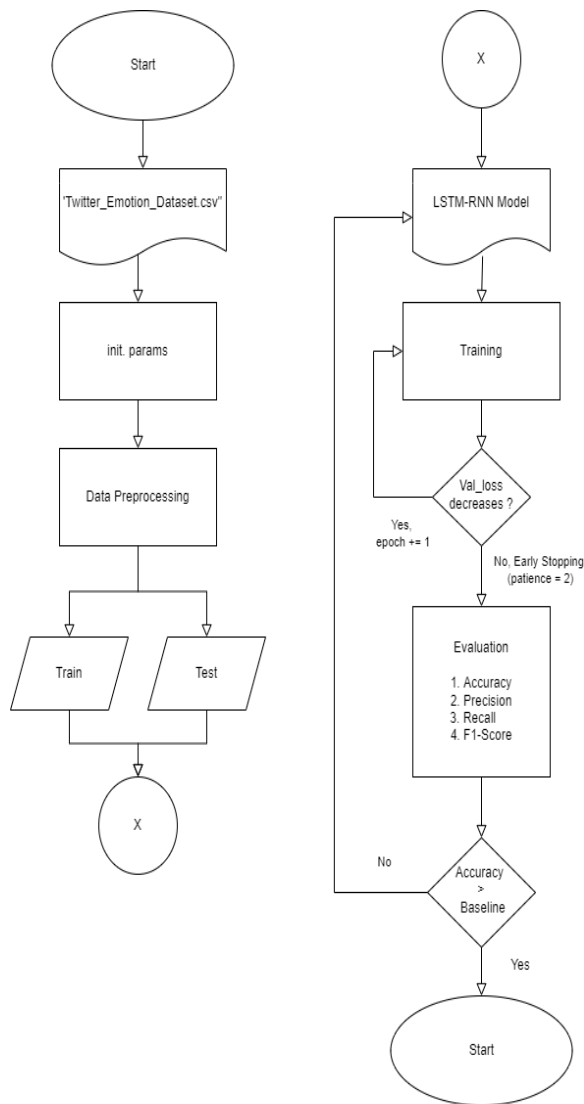
Pada penelitian[11] menggunakan Deep Learning LSTM-GloVe. Setelah memasuki lapisan input ke dalam urutan bilangan bulat, Tweet diberi token. Ditambahkan juga padding nol agar sesuai dengan panjang setiap tweet. Kemudian indeks embeddings mengambil indeks dari kamus GloVe. Parameter yang digunakan pada percobaan ini adalah: (1) Learning rate: 0,005 dan 0,001, (2) Dropout: 0,25 dan 0,5, (3) Optimizer: SGD dan Adam. Dengan 50 epoch dan 100 batch size, dari Penelitian ini didapatkan presisi, recall, dan F1-score masing-masing sebesar 33%, 38%, dan 35%.Keakuratan model adalah 46% [11]. Hasil yang didapatkan masih rendah. Penulis menyatakan bahwa hal itu disebabkan oleh underfitting. Namun penulis tidak menjelaskan alasan mereka membiarkan model tersebut menjadi underfit ketika proses pembelajaran sedang dilakukan.

Dalam penelitian[12], menggunakan metode CNN dengan parameter, seperti menambahkan 50% dropout, menerapkan regularisasi L2, dan menurunkan learning rate Adam Optimizer menjadi 0,0003 Hasil performansinya cukup tinggi, mengingat penelitian ini menggunakan dataset yang kecil. Presisi, recall, dan F1-score diperoleh masing-masing 90,1%, 90,3%, dan 90,2%, sedangkan akurasi adalah 89,8%. Berdasarkan penelitian [10][11][12], dengan dataset yang sama, metode CNN memimpin hasil F1-score sebesar 90,3%, dan Logistic Regression memiliki F1-score sebesar 69,73%, sedangkan metode LSTM-GloVe jauh lebih rendah dibandingkan Logistic Regression dan CNN.

Kontribusi dan tujuan dari penelitian kami adalah untuk mengembangkan dan mengoptimalkan penelitian [12] menggunakan metode yang berbeda yakni LSTM-RNN dan mendeteksi depresi pada tweet twitter, Penelitian[13] Identifikasi komentar kasar di media sosial Indonesia, menggabungkan RNN-LSTM mendapatkan hasil yang tinggi, dimana Presisi, recall, dan F1-score 94%, 95% dan 94%, sehingga penggabungan LSTM-RNN pada penelitian ini dilakukan untuk mendapatkan hasil akurasi yang memuaskan. Dataset yang digunakan sama dengan penelitian sebelumnya, nantinya dalam penelitian ini akan mengimplementasikan beberapa eksperimen pada arsitektur model (misalnya, penyetulan hyperparameter) untuk

mendapatkan hasil kinerja setinggi mungkin dan untuk menghindari overfits dan underfits.

METODE

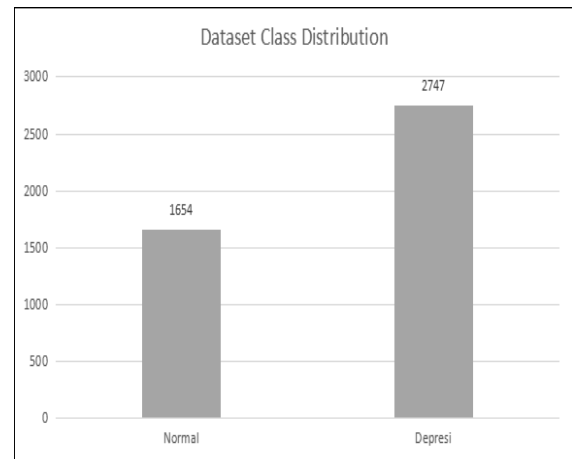


Gambar 1. System Flowchart

Pada penelitian ini, metode yang akan digunakan adalah kombinasi LSTM dan RNN. Gambar 1. merupakan flowchart alur dari sistem yang dibangun dalam penelitian ini. Sistem akan menerima inputan data dengan format CSV. Sebelum data digunakan, data terlebih dahulu dilakukan preprocessing. Lalu data yang sudah dilakukan preprocessing dibagi menjadi 70% data latih dan 30% data uji. Setelah dilakukan splitting data, selanjutnya membangun arsitektur model LSTM-RNN dan melakukan pelatihan terhadap model yang sudah dibangun. Model LSTM-RNN kemudian melakukan prediksi kelas

dari setiap data tweet didalam data uji dan melakukan evaluasi hasil kinerja model. Evaluasi yang digunakan adalah akurasi, presisi, recall, dan F1-Score.

1. Dataset



Gambar 2. Dataset Class Distribution

Sebelumnya dataset berisi 4.403 baris tweet yang memiliki lima label kelas yaitu marah, takut, gembira, cinta, dan sedih yang disediakan oleh [10] dalam format CSV. Setiap baris terdiri dari tweet dan label emosi. Pada penelitian ini dataset dibagi menjadi dua yaitu negative dan positif, Berdasarkan penelitian [14] mengatakan bahwa perasaan sedih, kecewa, mudah marah hal tersebut mencirikan bahwa orang tersebut sedang depresi.

Pada Tabel 1. Merupakan data yang sudah dilakukan pembagian menjadi 2 kelas. Untuk label sad, angry, dan fear masuk kedalam kategori negatif karena hal tersebut condong ke arah kelas depresi. Sedangkan untuk label happy dan love masuk ke dalam kelas normal karena emosi tersebut tidak berkaitan dengan depresi. Pada Gambar 2. data normal berjumlah 1.654 baris tweet dan data depresi berjumlah 2.747 baris tweet, Untuk memudahkan pengelolaan data kategori Depresi diberi label "1" sedangkan untuk kategori normal diberi label "0".

Tabel 1. Spesifikasi Dataset

Label	Tweet
1	soal jln jatibarupolisi tdk bs gerak gubernur emangny polisi tdk ikut pmbhasan jgn berpolitik pengaturan wilayahhak gubernur persoalan tn abang soal turun temurunpelikperlu kesabaran username username url

0 kepingin gudeg mbarek bu hj
amad foto dari google sengaja biar
temanteman membayangkannya
berbagi itu indah

Dilihat dari Gambar 2 menunjukkan bahwa distribusi data tidak seimbang, hal tersebut akan berpengaruh terhadap performa sistem, sehingga perlu adanya tindakan handling imbalance dataset, handling balancing dilakukan dengan melakukan balance class dengan teknik oversampling (meningkat data yang jumlahnya minoritas), oversampling dilakukan dengan mengambil data secara acak pada kategori normal (yang bernilai 0), setelah dilakukan balancing data, total data berjumlah 5.494 baris, dengan distribusi data normal dan depresi berjumlah 2.747.

2. Data Preprocessing

Untuk meningkatkan kualitas data, data dilakukan proses preprocessing terlebih dahulu sebelum masuk ke proses pelatihan. Dataset tersebut perlu melalui beberapa tahap data preprocessing sehingga dapat mendapatkan nilai hasil yang lebih bagus. Preprocessing data dalam Pembelajaran Mesin mengacu pada teknik penambangan data untuk membersihkan dan mengatur data mentah agar sesuai untuk membangun dan melatih model.

Dalam penelitian ini, kami menggunakan python untuk mengembangkan sistem dan Keras untuk mengevaluasi model pembelajaran. Keras merupakan API pembelajaran mendalam yang konsisten dan sederhana, berjalan di atas TensorFlow, dirancang untuk meminimalkan jumlah tindakan pengguna yang diperlukan untuk kasus penggunaan umum, dan menyediakan pesan kesalahan yang jelas dan dapat ditindaklanjuti[9][15].

Pada bagian preprocessing, Tokenizer dari Keras akan diterapkan untuk mendapatkan indeks kata dengan cara tokenizing kata-kata di dalam dokumen teks, juga untuk mengubah setiap teks menjadi urutan bilangan bulat. Urutan akan digunakan untuk komputasi pada jaringan saraf karena tidak dirancang untuk memahami input non-numerik.

2.1. Case-Folding

Pada tahap ini mengubah keseluruhan teks menjadi huruf kecil, Kemudian teks dilakukan pembersihan terhadap nomer, url, white space, serta tanda baca yang dinilai tidak memiliki pengaruh atau memberikan pengaruh

buruk terhadap pemrosesan text. Proses Case folding dapat dilihat pada Tabel 2.

Tabel 2. Hasil Proses Case-Folding

Tweet	Case-Folding
Soal jln Jatibaru, polisi tdk bs GERTAK gubernur .Emangny polisi tdk ikut pmbhasan? Jgn berpolitik. Pengaturan wilayah, hak gubernur. Persoalan Tn Abang soal turun temurun. Pelik. Perlu kesabaran. [USERNAME] [USERNAME] [URL]	soal jln jatibarupolisi tdk bs gertak gubernur emangny polisi tdk ikut pmbhasan jgn berpolitik pengaturan wilayahhak gubernur persoalan tn abang soal turun temurunpelikperlu kesabaran username username url
Sesama cewe lho (kayaknya), harusnya bisa lebih rasain lah yang harus sibuk jaga diri, rasain sakitnya haid, dan paniknya pulang malem sendirian. Gimana orang asing? Wajarlah banyak korban yang takut curhat, bukan dibela malah dihujat.	sesama cewe lho kayaknya harusnya bisa lebih rasain lah yang harus sibuk jaga diri rasain sakitnya haid dan paniknya pulang malem sendirian gimana orang asing wajarlah banyak korban yang takut curhat bukan dibela malah dihujat

2.2. Tokenizing

Tahapan tokenizing merupakan proses pemecahan kalimat menjadi potongan-potongan kata, tanda baca, dan ungkapan yang bermakna lainnya berdasarkan ketentuan bahasa. Dengan menggunakan Natural Language Toolkit (NLTK). Proses tokenizing dapat dilihat pada Tabel 3.

Tabel 3. Proses Tokenizing

Tweet	Tokenizing
via nulis dude dude kan bukan pemain bola twitter receh ig receh kelen semua emang kembalian mini market recehhhhh	['via', 'nulis', 'dude', 'dude', 'kan', 'bukan', 'pemain', 'bola', 'twitter', 'receh', 'ig', 'receh', 'kelen', 'semua', 'emang', 'kembalian', 'mini', 'market', 'recehhhhh']

2.3. Stop Words Removal

Stop words Removal adalah proses filtering menggunakan stopwords removal untuk pemilihan kata-kata penting dengan menghilangkan kata singkatan dan kata tidak baku (contoh: yg, dgn, kalo, dsb) serta kata-kata umum (dan, dengan, adalah, akan, dsb) yang tidak memberi arti dalam bahasa Indonesia. proses ini menggunakan library Natural Language Toolkit (NLTK). Tabel 4. Merupakan proses Stop Words Removal.

Tabel 4. Proses Stop Words Removal

Tweet	Stop Words Removal
['karena', 'makan', 'bersama', 'sambil', 'berbincang', 'dengan', 'orang', 'tersayang', 'itu', 'terasa', 'nikmat', 'by', 'me', 'taichangoreng']	['makan', 'berbincang', 'orang', 'tersayang', 'nikmat', 'by', 'me', 'taichangoreng']
['sharing', 'pengalaman', 'aja', 'kemarin', 'jam', 'batalin', 'tiket', 'di', 'stasiun', 'pasar', 'senen', 'lancar', 'antrian', 'tidak', 'terlalu', 'rame', 'menitan', 'dan', 'beress', 'semua', 'mungkin', 'bisa', 'dicoba', 'twips', 'di', 'jamjam', 'segitu', 'cc', 'username']	['sharing', 'pengalaman', 'kemarin', 'jam', 'batalin', 'tiket', 'stasiun', 'pasar', 'senen', 'lancar', 'antrian', 'rame', 'menitan', 'beress', 'dicoba', 'twips', 'jamjam', 'segitu', 'cc', 'username']

2.4. Slang Words Removal

Slang words Removal adalah proses menghilangkan kata-kata bahasa tidak baku, serta ada beberapa bahasa yang disingkat-singkat, atau adanya penggunaan bahasa gaul yang tidak sesuai dengan kaidah bahasa Indonesia. Sehingga, diperlukan proses normalisasi untuk mengubah kata-kata atau

karena makan bersama sambil berbincang dengan orang tersayang itu terasa nikmat by me taichangoreng	['karena', 'makan', 'bersama', 'sambil', 'berbincang', 'dengan', 'orang', 'tersayang', 'itu', 'terasa', 'nikmat', 'by', 'me', 'taichangoreng']
---	--

bahasa yang tidak sesuai kaidah bahasa Indonesia menjadi bahasa yang dapat dimengerti oleh khalayak umum. Proses slang word removal dapat dilihat pada Tabel 5.

2.5. Stemming

Stemming adalah proses mengidentifikasi akar/batang suatu kata dan menghilangkan kata imbuhan seperti awalan, akhiran, dengan menggunakan fungsi Stemmer Factory yang terdapat pada library Sastrawi. Hasil stemming dapat dilihat pada Tabel 6.

Tabel 5. Proses Slang Words Removal

Tweet	Slang Words Removal
['selamaaaat', 'paaaagi', 'menjelang', 'siang', 'sayang', 'kuuu', 'semangat', 'jaga', 'kesehatan', 'sayang', 'kamui', 'love', 'bby', 'username']	['selamat', 'ulang', 'dek', 'zein', 'tahun', 'semoga', 'umur', 'pintar', 'cerdas', 'anak', 'sholih', 'berbakti', 'orang', 'tua', 'nusa', 'bangsa', 'uniform resource locator']
['menit', 'menit', 'gw', 'udah', 'jatuh', 'cinta', 'lu', 'salah', 'sapa', 'coba']	['menit', 'menit', 'gue', 'sudah', 'jatuh', 'cinta', 'kamu', 'salah', 'sapa', 'coba']

Tabel 6. Indonesian Word Stemming

Word	Stemming
Keuangan	Uang
Pertimbangan	Timbang
Indahnya	Indah
Andalan	Andal

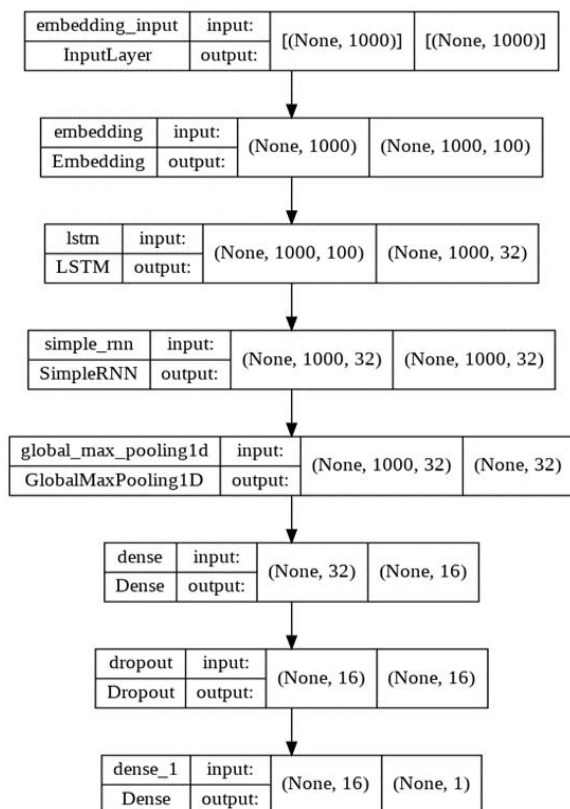
3. Arsitektur Model

Pada penelitian ini, metode yang akan digunakan adalah LSTM-RNN. Gambar 4. merupakan arsitektur model yang digunakan dalam penelitian ini.

3.1. Embedding Layer

Embedding Layer bekerja dengan cara yang sama sebagai perkalian matriks sederhana dengan mengubah kata menjadi penyisipan kata yang sesuai atau mengubah indeks menjadi vektor padat dengan ukuran tetap. Penyisipan kata dianggap sebagai fitur tekstual sehingga dapat dihitung sebagai langkah preprocessing dalam tugas NLP [16].

Lapisan ini digunakan sebagai lapisan pertama dalam model. Data input akan diisi untuk setiap kalimat untuk membuat ukuran urutan yang tetap. Lapisan embedding akan menerima urutan ini dan membuat matriks embedding dalam ukuran $m \times n$, di mana m adalah dimensi output. Terdiri dari korelasi setiap indeks kata di setiap tweet. Dalam model LSTM-RNN kami, menggunakan dimensi embedding 100, lapisan embedding akan belajar bersamaan dengan model, dan tidak menggunakan pre-trained embedding seperti Word2Vec dan fastText.



Gambar 4. Model Architecture

3.2. LSTM

LSTM berisikan unit khusus yang disebut blok memori pada recurrent hidden layer. Blok memori berisi sel memori dengan koneksi sendiri yang menyimpan status temporal

jaringan di samping unit multiplikasi khusus yang disebut gerbang untuk mengontrol aliran informasi. Setiap blok memori dalam arsitektur asli berisi gerbang input dan gerbang output[17]. Keunggulan LSTM adalah kemampuan mengingat dari sekuens long term (ukuran data) yang sulit dicapai dengan teknik fitur tradisional[18].

3.3. RNN

Recurrent neural networks (RNNs) berisikan koneksi siklik yang menjadikannya alat yang lebih kuat untuk memodelkan data urutan seperti itu daripada jaringan saraf umpan maju. RNN telah menunjukkan keberhasilan besar dalam pelabelan urutan dan tugas prediksi seperti pengenalan tulisan tangan dan pemodelan bahasa[17].

3.4 Kombinasi LSTM-RNN

Kombinasi RNN dan LSTM digunakan untuk mengembangkan model yang optimal. Data latih tersebut dimasukkan ke dalam model klasifikasi LSTM dan RNN sebagai input data. Lapisan input didefinisikan sebagai lapisan embedding, dimana pada lapisan ini terdapat proses pencarian kosakata yang dikodekan dengan array bilangan bulat dan vektor penyematan untuk setiap indeks kata. Dimensi yang dihasilkan adalah ukuran batch, urutan, dan ukuran embedding. Dimensi penyematan mengatur jumlah fitur untuk kata yaitu, jumlah unit tersembunyi. Hasil embedding berupa matriks dengan panjang kamus dan dimensi embedding.

Kemudian, memasukkan blok memori LSTM, LSTM berisi unit khusus disebut blok memori di lapisan tersembunyi berulang dari modul tensorflow.

Selanjutnya menambahkan Layer RNN dengan SimpleRNN memasukkan blok memori. Kemudian, ditambahkan lapisan dense untuk membuat lapisan NN kompleksitas dan dropout digunakan untuk menangani masalah overfitting [18].

Terakhir keluaran data yang direpresentasikan dengan nilai biner (0 dan 1). Nilai-nilai ini adalah probabilitas, 0 adalah untuk data yang normal dan 1 untuk data yang depresi. Dan Label ini akan digunakan untuk mengklasifikasikan kalimat depresi.

3.5. Pooling Layer

Lapisan pooling adalah lapisan yang mengurangi input secara spasial dengan menggunakan operasi downsampling. Dalam penelitian ini, metode pooling yang digunakan adalah global max pooling. Global max pooling

melakukan operasi hitung nilai max tunggal untuk setiap saluran input. Global max pooling mengambil tensor 2 dimensi ukuran (ukuran input) x (saluran input) dan menghitung maksimum semua nilai (ukuran input) untuk masing-masing (saluran input). lapisan pooling global menurunkan sampel seluruh peta fitur menjadi satu nilai. Ini akan sama dengan mengatur ukuran penyatuan ke ukuran peta fitur input.

3.6. Dense

Lapisan dense merupakan lapisan jaringan saraf yang terhubung secara mendalam, dimana setiap neuron pada lapisan padat menerima input dari semua neuron pada lapisan sebelumnya. lapisan dense melakukan perkalian matriks-vektor. Output yang dihasilkan oleh lapisan dense adalah vektor dimensi 'm'. Lapisan dense, paling umum digunakan dalam model.

3.7. Dropout

Untuk menghindari peluang overfit saat melatih model, disebabkan oleh Dataset yang kecil. Hal tersebut mempengaruhi kinerja model ketika proses pelatihan telah dilakukan. Untuk mengurangi overfitting dari setiap jaringan dan rata-rata prediksi model. Dropout juga diterapkan setelah layer LSTM untuk mengontrol overfitting[18]. Dengan menggunakan dropout untuk mengatur model ukuran tetap dengan merata-ratakan prediksi hiperparameter yang digunakan dalam lapisan tertentu dengan memberi bobot pada setiap pengaturan probabilitas posterior yang diberikan data pelatihan[19].

Dropout dapat mengganggu proses pelatihan dengan memaksa lapisan untuk menjatuhkan node mereka. Nilai dropout adalah antara 0 dan 1. Dengan asumsi itu ditentukan ke 0,5, itu akan memaksa lapisan untuk menjatuhkan 50% dari node mereka untuk mengurangi overfitting.

3.8. Reduce LR On Plateau

Reduce LR On Plateau adalah teknik penjadwalan yang memantau kuantitas dan meluruhkan kecepatan belajar ketika kuantitas berhenti meningkat. Peningkatan kuantitas didasarkan pada apakah itu meningkat atau menurun dengan jumlah minimum tertentu. Jumlah minimum adalah ambang batas.

3.9. Adam Optimizer

Adam Optimizer adalah algoritma pengoptimalan laju pembelajaran stokastik adaptif yang dirancang untuk jaringan saraf

dalam dengan menghitung laju pembelajaran individu untuk parameter yang berbeda. optimizer Adam yaitu merupakan perkembangan dari algoritma Stochastic Gradient Descent (SGD) klasik dimana bobot network telah diperbarui[20]. Algoritma ini pertama kali dikenalkan oleh Diederik Kingma[21]. Dengan menerapkan konsep optimasi stokastik yang sama. Optimasi stokastik merupakan proses mengoptimalkan fungsi tujuan. Adam Optimizer menangani sparse gradients pada dataset yang bising secara efisien dengan menggunakan jumlah memori yang lebih kecil dan dapat menjadi keuntungan untuk penelitian ini

4. Model Testing

Untuk memastikan kinerja model yang baik diperlukan uji dengan melakukan prediksi. Model yang dilatih sebelumnya akan memprediksi label dari data uji. Dalam penelitian ini, metode yang digunakan untuk mengevaluasi prediksi adalah akurasi, presisi, recall, dan FMeasure. Metode-metode ini diukur berdasarkan matriks konfusi yang disajikan pada Tabel 7.

Tabel 7. Confusion Matrix

	Actual Class (+)	Actual Class (-)
Predicted Class (+)	TP	FP
Predicted Class (-)	FN	TN

Akurasi sering digunakan dalam penelitian yang berkaitan dengan klasifikasi biner dan/atau klasifikasi multi kelas[22]. Akurasi merupakan jumlah titik data yang diprediksi dengan benar dari semua titik data. Akurasi dihitung dengan membagi jumlah prediksi yang benar dengan jumlah total prediksi, Persamaan (1) merupakan rumus untuk mendapatkan nilai akurasi.

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

Pada persamaan (1), TP direpresentasikan sebagai True Positives, di mana model memprediksi kelas positif dengan benar. TN berarti True Negatives, yang berarti model telah memprediksi kelas negatif dengan benar. FP dan FN adalah singkatan dari False Positives dan False Negatives. menunjukkan model telah salah memprediksi kelas positif dan negatif.

Presisi adalah seberapa sering memprediksi kelas positif dengan benar. Untuk menghitung presisi dapat menggunakan Persamaan (2).

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

Recall adalah probabilitas label positif berlabel positif. Untuk menghitung recall dapat menggunakan Persamaan (3).

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

F-Measure memberikan skor tunggal yang menyeimbangkan perhatian terhadap presisi dan daya ingat dalam satu angka. perbandingan jumlah kelas positif yang diprediksi benar dengan jumlah semua sampel yang seharusnya diprediksi positif. Metode ini umumnya digunakan untuk mengevaluasi berbagai jenis model pembelajaran mesin. F-Measure dapat dihitung dengan menggunakan Persamaan (4).

$$F - Measure = \frac{2PR}{P + R} \quad (4)$$

F-Measure dihasilkan dengan memiliki trade-off antara Precision dan Recall. Ketika presisi dan recall sama atau memiliki nilai yang sedikit berbeda, F-Measure pasti akan memiliki hasil yang bagus.

HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan sebuah sistem yang mampu melakukan klasifikasi berdasarkan kelasnya pada sistem klasifikasi depresi. Selanjutnya dilakukan pengujian kinerja sistem, Pengujian dilakukan dengan memasukkan index di dalam dataset sebagai sampel yang kemudian sistem memprediksi depresi apakah masuk kedalam kelas depresi atau normal.

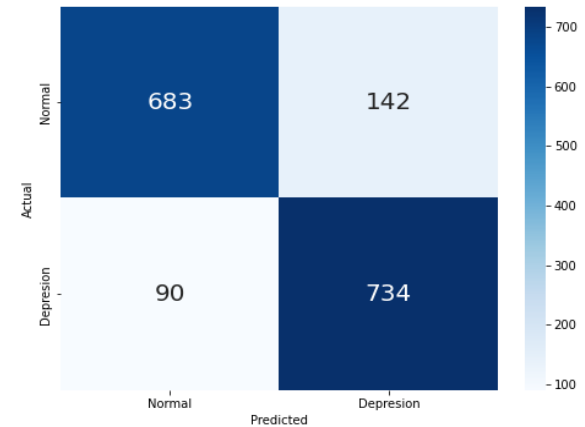
1. Training Process

Berdasarkan Gambar 4, lapisan input menerima urutan bilangan bulat dari data yang telah diproses sebelumnya. Data terdiri dari 10520 kosakata unik dan maksimal 1000 kata untuk seluruh dokumen teks. Lapisan embedding mengambil setiap kosakata unik sebagai dimensi input, 100 dimensi input dan output, dan 1000 panjang input maksimum. Karena lapisan embedding adalah lapisan pertama dari model, maka perlu memiliki lebih banyak data dari lapisan input untuk matriks

embedding dalam hal proses pembelajaran yang lebih baik.

Untuk mengoptimalkan model dilakukan penambahan layer arsitektur berupa: (1) GlobalMaxPool1D, (2) Dense Layer 16 dengan activation 'relu', (3) Dropout layer sebesar 0.5, diakhiri dengan (4) Dense layer dengan output terdiri atas 2 kelas dengan activation 'sigmoid', serta optimizer 'adam'.

2. Text Depresi Classification



Gambar 5. Confusion Matrix

Prediksi dilakukan dengan mengelompokkan tweet ke dalam 2 kelas. Confusion matrix pada Gambar 5. menunjukkan tweet yang diprediksi dengan benar di antara kelas mereka. 683 tweet diklasifikasikan dengan benar sebagai Normal, 734 sebagai depresi, 142 tweet normal diklasifikasikan salah, 90 tweet depresi diklasifikasikan salah.

	precision	recall	f1-score	support
0	0.88	0.83	0.85	825
1	0.84	0.89	0.86	824
accuracy			0.86	1649
macro avg	0.86	0.86	0.86	1649
weighted avg	0.86	0.86	0.86	1649

Gambar 6. Classification Report

Gambar 6, menunjukkan hasil evaluasi masing-masing kelas prediksi. Hasil ini didasarkan pada matriks konfusi pada Gambar 5. Metode evaluasi yang digunakan adalah akurasi, presisi, recall, dan F1-score. Pada Gambar 6, skor presisi tertinggi adalah kelas Normal, sebesar 88%. Sedangkan kelas depresi memiliki skor recall tertinggi sebesar 89%. Kelas normal memiliki rata-rata hasil terendah dari semuanya. Hal ini menunjukkan bahwa sistem mengalami lebih banyak kesulitan dalam mengklasifikasikan tweet ke kelas ini. Mungkin

karena tren tweet berlabel normal terlalu rumit untuk dipelajari modelnya. Namun demikian, hasil tersebut tidak signifikan dalam hal interval skor, sehingga model akan tetap mendapatkan hasil kinerja keseluruhan yang tinggi. Pada Tabel 8, dibandingkan hasil akurasi dengan penelitian-penelitian sebelumnya.

Tabel 8. Perbandingan Hasil Model

Metode	Akurasi	Presisi	Recall	F1-Score
LSTM-Glove	46%	33%	38%	35%
CNN	89,8%	90,1%	90,3 %	90,2 %
LSTM-RNN	86%	86%	86%	86%

Pada Tabel 9. Kami melakukan pengujian terhadap model, Pengujian dilakukan dengan memasukkan index di dalam dataset sebagai sampel yang kemudian sistem memprediksi dan akan menampilkan output prediksi berupa depresi atau normal.

Tabel 9. Hasil Prediksi

Tweet	Prediksi
'ku' 'khawatir' 'di dunia' 'khawatir'	Depresi
'transj' 'gue' 'diri' 'baik' 'hati' 'kasih' 'duduk' 'tua' 'tetapi' 'tua' 'eh' 'cowok' 'sopan' 'duduk' 'duduk' 00_V0	Normal
'doa' 00_V 'lulus' 'tes' 'besok' 'surabaya' 'bu' 'jujur' 'isi' 'materi' 'forum' 'level' 'dokter' 'benar' 'takut' 'takut' 'salah' 'maksud'	Depresi
'diam' 'bagus' 'ya sudah' 'diam' 'deh' 'ingin tahu' 'deh' 00_V 'haha'	Normal

KESIMPULAN

Dalam penelitian ini, metode LSTM-RNN yang diimplementasikan untuk mengidentifikasi emosi dengan mengklasifikasikan tweet dalam bahasa Indonesia dengan mengatur kombinasi parameter, seperti menggunakan 50% dropout, dan menurunkan learning rate Adam Optimizer. Hasil performansinya cukup tinggi, mengingat

penelitian ini menggunakan dataset yang kecil. Presisi, recall, dan F1-score diperoleh masing-masing 86%, 86%, dan 86%, sedangkan akurasi adalah 86%. Perlu disebutkan bahwa kombinasi parameter memainkan peran besar dalam mengoptimalkan model.

Sistem deteksi ujaran depresi diharapkan dapat membantu menganalisa depresi masyarakat di dunia maya, yang dilihat berdasarkan hasil keluaran sistem yang selanjutnya dapat dianalisis lebih lanjut. Saran untuk penelitian selanjutnya adalah dapat mengumpulkan data depresi yang lebih banyak untuk setiap tingkatan depresi agar menghasilkan akurasi yang lebih baik.

UCAPAN TERIMAKASIH

Ucapan terima kasih disampaikan kepada semua pihak yang telah berperan dalam penelitian ini sehingga penelitian dapat terlaksana dengan baik.

REFERENSI

- [1] M. Abdullah and S. Shaikh, "TeamUNCC at SemEval-2018 Task 1: Emotion Detection in English and Arabic Tweets using Deep Learning," NAACL HLT 2018 - Int. Work. Semant. Eval. SemEval 2018 - Proc. 12th Work., pp. 350–357, 2018, doi: 10.18653/v1/s18-1053.
- [2] Simon Kemp, "Digital 2022: Indonesia-DataReportal," Global Digital Insights, 2022.
- [3] A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of Social Media," Bus. Horiz., vol. 53, no. 1, 2010, doi: 10.1016/j.bushor.2009.09.003.
- [4] S. Stieglitz, M. Mirbabaie, B. Ross, and C. Neuberger, "Social media analytics – Challenges in topic discovery, data collection, and data preparation," Int. J. Inf. Manage., vol. 39, no. December 2017, pp. 156–168, 2018, doi: 10.1016/j.ijinfomgt.2017.12.002.
- [5] K. Pranav, "Neural Network Based System to Detect Depression in Twitter Users via Sentiment Analysis," Int. Res. J. Eng. Technol., pp. 1449–1451, 2018, [Online]. Available: www.irjet.net
- [6] M. K. Neighbor, A. P. Tirtopangarsa, W. Maharani, T. Informasi, and U. Telkom, "Sentiment Analysis of Depression Detection on Twitter Social Media Users Using the K-Nearest Neighbor Method," J. SEMNASIF, pp. 247–258, 2021.

- [7] B. Y. Ziwei and H. N. Chua, "An application for classifying depression in tweets," 2019. doi: 10.1145/3366650.3366653.
- [8] A. Chatterjee, U. Gupta, M. K. Chinnakotla, R. Srikanth, M. Galley, and P. Agrawal, "Understanding Emotions in Text Using Deep Learning and Big Data," *Comput. Human Behav.*, vol. 93, 2019, doi: 10.1016/j.chb.2018.12.029.
- [9] M. Abadi et al., "Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv 2016," arXiv preprint arXiv:1603.04467. 2019.
- [10] M. S. Saputri, R. Mahendra, and M. Adriani, "Emotion Classification on Indonesian Twitter Dataset," 2019. doi: 10.1109/IALP.2018.8629262.
- [11] A. D. L. B. Harijanto, and F. Rahutomo, "Implementasi Deep Learning Untuk Deteksi Ekspresi Emosi Pada Twitter," pp. 3–6, 2020.
- [12] N. Hilmiaji, K. M. Lhaksana, and M. D. Purbolaksono, "Identifying Emotion on Indonesian Tweets using Convolutional Neural Networks," no. 10, pp. 584–593, 2021.
- [13] T. I. Sari, Z. N. Ardilla, N. Hayatin, and R. Maskat, "Abusive comment identification on Indonesian social media data using hybrid deep learning," *IAES Int. J. Artif. Intell.*, vol. 11, no. 3, pp. 895–904, 2022, doi: 10.11591/ijai.v11.i3.pp895-904.
- [14] A. Dirgayunita, "Depresi: Ciri, Penyebab dan Penangannya," *J. An-Nafs Kaji. Penelit. Psikol.*, vol. 1, no. 1, 2016, doi: 10.33367/psi.v1i1.235.
- [15] F. Chollet and O., "Keras: the Python deep learning API," Keras: the Python deep learning API, 2020.
- [16] Y. Li and T. Yang, "Word Embedding for Understanding Natural Language: A Survey," 2018. doi: 10.1007/978-3-319-53817-4_4.
- [17] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," 2014. doi: 10.21437/interspeech.2014-80.
- [18] M. Rizki, S. Basuki, and Y. Azhar, "Implementasi Deep Learning Menggunakan Arsitektur Long Short Term Memory(LSTM) Untuk Prediksi Curah Hujan Kota Malang," *J. Repos.*, vol. 2, no. 3, 2020, doi: 10.22219/repositor.v2i3.470.
- [19] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, 2014.
- [20] N. Rochmawati, H. B. Hidayati, Y. Yamasari, H. P. A. Tjahyaningtjas, W. Yustanti, and A. Prihanto, "Analisa Learning Rate dan Batch Size pada Klasifikasi Covid Menggunakan Deep Learning dengan Optimizer Adam," *J. Inf. Eng. Educ. Technol.*, vol. 5, no. 2, 2021, doi: 10.26740/jieet.v5n2.p44-48.
- [21] "Keras: Introduction to the Adam Optimization Algorithm," OnnoCenterWiki. https://lms.onnocenter.or.id/wiki/index.php/Keras:_Introduction_to%0A_the_Adam_Optimization_Algorithm (accessed Nov. 09, 2021)
- [22] M. Hossin, M. N. Sulaiman, A. Mustapha, N. Mustapha, and R. W. Rahmat, "A hybrid evaluation metric for optimizing classifier," 2011. doi: 10.1109/DMO.2011.5976522.