DETEKSI KECEMASAN DI TWITTER MENGGUNAKAN FITUR WORD EMBEDDING BERT DAN METODE BIDIRECTIONAL-LONG SHORT TERM MEMORY

Angga Prawira, Yulison Herry Chrisnanto, Ade Kania Ningsih
Progam Studi Teknik Informatika, Fakultas Sains dan Informatika, Universitas Jenderal Achmad Yani
Jl terusan jenderal sudirman, Cimahi, 4052, Indonesia
y.chrisnanto@lecturer.unjani.ac.id

ABSTRAK

Adanya stigma di mana pengidap gangguan jiwa adalah orang yang sakit jiwa membuat orang-orang yang mengalami kecemasan berlebih malu untuk berobat dan lebih memilih untuk mengeluarkan keluh kesah, perasaan mereka di sosial media oleh karena itu dibutuhkan sebuah model perangkat lunak yang dapat mendeteksi unggahan kecemasan berlebih. Pada penelitian sebelumnya sudah mendeteksi kecemasan berlebih di media sosial twitter menggunakan metode BI-LSTM dengan pembobotan kata menggunakan TF-IDF dan mendapatkan akurasi sebesar 94.12%. Metode BI-LSTM sendiri membutuhkan data set yang sangat besar agar model tidak over-fitting. Oleh karena itu, proses pembobotan kata dilakukan menggunakan BERT karena selain memberikan pembobotan kata, BERT juga memberikan representasi numerik sehingga data set lebih beragam dan bisa membuat model BI-LSTM lebih banyak belajar dari representasi numerik tersebut hasil dari penelitian ini didapatkan akurasi sebesar 99.44% dan loss 0.0278 dapat disimpulkan bahwa bert dapat meningkatkan akurasi dari metode Bi-LSTM sehingga metode ini cocok untuk deteksi kecemasan berlebih.

Kata kunci: *Kecemasan Berlebih; Bert; BI-LSTM*
\

1. PENDAHULUAN

Kecemasan berlebih merupakan salah satu gangguan mental di mana penderita akan merasa cemas terhadap suatu hal untuk waktu yang lama dan terkadang alasan cemasnya tidak jelas. Kecemasan berlebih ini bisa datang bersamaan dengan gangguan mental lainnya yaitu depresi di mana seseorang yang mengalami depresi akan merasakan perasaan sedih dan kehilangan minat terhadap perilaku sehari-hari. Menurut *World Health Organization* atau WHO orang yang mengalami depresi dan kecemasan berlebih meningkat 25% setelah pandemi covid 19. Adanya stigma dimasyarakat yang menganggap jika penderita gangguan mental mengidap gangguan jiwa[8] juga membuat penderita gangguan mental takut untuk berobat atau mengunjungi profesional untuk berkonsultasi. Gangguan mental jika dibiarkan dalam waktu yang lama akan berdampak pada penyembuhan yang akan semakin lama jika dibiarkan, mereka lebih memilih untuk mencurahkan isi hati atau perasaan mereka ke media sosial untuk mencari dukungan moral.

Sosial media merupakan sebuah platform yang sangat besar di mana setiap pengguna bebas berekspresi, mengunggah sesuatu, mengeluarkan perasaan dan mengeluarkan isi pikirannya[1]. Pengguna aktif sosial media di Indonesia berjumlah 167 juta atau 60.4% dari total penduduk Indonesia berdasarkan laporan *we are social* penduduk Indonesia menempati salah satu negara dengan jumlah orang yang mengakses internet terbanyak. Menurut katadata jumlah pengguna aktif Twitter atau sekarang X dari Indonesia berjumlah 24 juta atau peringkat 5 pengguna terbanyak .

Unggahan isi tweet yang menggambarkan depresi di sosial media dapat di deteksi menggunakan beberapa algoritma di antaranya ada menggunakan metode XGBoost menggunakan fitur *Fast Text embeddings* menghasilkan akurasi sebesar 71.5% untuk label ada 2 yaitu depresi dan bukan depresi[1]

Selain itu ada juga menggunakan metode LSTM di mana data set berasal dari reddit[2] mendapatkan akurasi sebesar 94.88%. Lalu ada juga penelitian menggunakan 2 metode yaitu LSTM dan CNN mendapatkan akurasi 94.03% [4] di mana data set menggunakan bahasa bangladesh. Lalu ada juga penelitian menggunakan metode Bi-LSTM mendapatkan akurasi 94.12% data set berbahasa Indonesia dan data set didapatkan dari media sosial twiter menggunakan pembobotan kata *Term Frequency-Term Inverse Document Frequency(TF-IDF)* di mana pembobotan kata berdasarkan kemunculan kata tersebut.

Bidirectional-Long Short Term Memory(Bi-Lstm) merupakan salah satu metode dari *Recurent Neura Network(RNN)* yang di mana merupakan pengembangan dari Long Short Term Memory di mana dalam pembelajaran mesin dari RNN memiliki masalah dalam dependensi jangka panjang sementara BI-LSTM sendiri merupakan pengembangan dari LSTM di mana BI-LSTM sendiri bekerja dengan membaca informasi secara dua arah dari sebuah kalimat jika konteks data setnya adalah kalimat BI-LSTM membutuhkan data set yang lumayan besar agar model tidak *over-fitting[8]* .

Word embedding merupakan salah satu metode untuk pengolahan kata di bidang pemrosesan bahasa alami atau NLP di mana word embdding memberikan

mengubah kata menjadi numerik agar dapat diolah oleh komputer. *Bidirectional Encoder Representation (BERT)*merupakan salah satu metode *word embedding* dinamis di mana selain memberikan pembobotan kata BERT juga memberikan representasi numerik dari yaitu hubungan kata dengan kata lain nya. BERT sendiri dapat digunakan untuk mendeteksi teks depresi atau bukan mendapatkan akurasi sebesar 71%[7] lalu BERT juga pernah digunakan untuk mendeteksi sentimen analisis optimis atau pesimis setelah covid 19 [3]. Penelitian ini dikerjakan menggunakan dua metode yaitu BERT untuk pembobotan kata dan mendapatkan representasi numerik dari data set dan menggunakan BI-LSTM untuk proses klasifikasi dengan label kecemasan berlebih dan bukan kecemasan berlebih.

2. TINJAUAN PUSTAKA

2.1. Word Embedding

Merupakan salah satu teknik analisis teks dalam ilmu pengolahan bahasa alami di mana *word embedding* bekerja dengan cara mengubah teks menjadi vektor numerik [4] yang dapat digunakan untuk mengelompokkan teks berdasarkan pola tertentu. Klasifikasi teks ini dapat digunakan juga untuk mendeteksi teks.

Salah satu keuntungan menggunakan *Word embedding* adalah *word embedding* melihat konteks hubungan semantik antar kata sebagai contoh raja dan ratu akan memiliki nilai vektor yang terkait [9]. Keuntungan tersebut merupakan salah satu kelemahan dari *one hot encoding* di mana *one hot encoding* tidak dapat mengasosiasikan hubungan semantik antar kata.

2.2. Bert

Merupakan salah satu metode dari word embedding di mana berfungsi untuk mengubah representasi teks menjadi vektor numerik menggunakan metode transformers di mana transformers termasuk contextualized word embedding. Bert memakai Transformer untuk mempelajari hubungan kontekstual antar kata[13]. Berbeda dengan trasnformer biasa bert tidak bekerja sequential di mana membaca kalimat dari kiri ke kanan bert melihat keseluruhan hubungan antar kata dari kalimat tersebut[13].

Bert sendiri memiliki beberapa fitur seperti mask masked languaged di mana akan ada 15% ukuran token dengan mask lalu bert akan berusaha menemukan nilai mask itu sesuai dengan nilai aslinya[11].

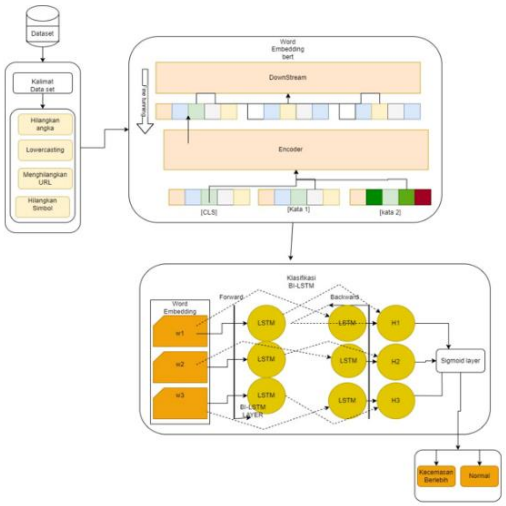
2.3. BI-LSTM

Long Sort – Term Memory (LSTM) merupakan arsitektur pengembangan dari Recurent Neural Network (RNN) digunakan untuk mengatasi masalah *vanishing* gradient di mana kemiringan fungsi kerugian menurun secara eksponensial di saat memproses data *sequential* yang panjang [8]. Masalah itulah yang menyebabkan RNN tidak bisa mendapatkan *long term despedency* yang berakibat pada menurun nya peforma prediksi.

Salah satu kelemahan LSTM adalah tidak cukup memperhitungkan kata yang terletak di akhir karna hanya bekerja satu arah yaitu dari awal ke akhir saja [8] maka dengan adanya kekurangan itu disarankan untuk menggunakan BI-LSTM yaitu LSTM yang sudah dikembangkan di mana dapat bekerja dua arah yaitu dari awal ke akhir dan dari akhir ke awal.

3. METODOLOGI PENELITIAN

Model deteksi unggahan berbentuk kalimat yang di klasifikasikan sebagai kalimat kecemasan berlebih atau bukan kecemasan berlebih terdiri dari tiga tahapan yaitu pre-processing, word embedding bert dan klasifikasi menggunakan bi-lstm seperti pada gambar 1.



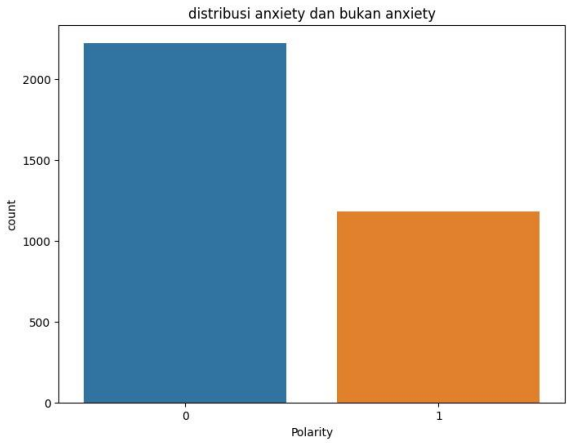
Gambar 1. Metodologi penelitian

3.1. Data set

Data set di dapatkan dari kaggle di mana memiliki 2 label yaitu kecemasan berlebih dan bukan kecemasan berlebih. Untuk jumlah data set sendiri terdapat 2224 berlabel bukan kecemasan berlebih dan 1180 data berlabel kecemasan berlebih seperti pada gambar 2 di mana 0 berarti label bukan kecemasan berlebih sementara label 1 berarti kecemasan berlebih dan seperti pada tabel 1 merupakan 5 sampel dari data set.

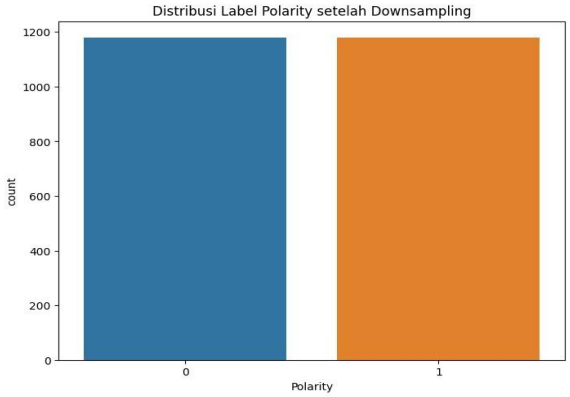
Tabel 1. isi dari data set

Tweet	Polarity
udaa 2 harian tidur gue ga nyenyak bgt, kaya gelisah gt. kenapa yaaa :([]	1
Ini gua kenapa ya...Udah mood berantakan trus tuh jantungnya kayak detaknya kuat banget... gelisah banget.	1
Gue sudah anxiety di line anak sefakultas ternyata hanya minta foto buat sg akun angkatan asem	0
Anxiety kumat diminggu sore menjelang senin kerjaaaa	0



Gambar 2. Pengelompokan data set berdasarkan label

Berdasarkan pengelompokan data di atas terdapat perbedaan dalam jumlah data dari label 0 dan 1 lebih banyak label 0 atau bukan kecemasan berlebih oleh karena itu diperlukan downsampling untuk mengurangi jumlah label 0 agar jumlah distribusi data set sama seperti gambar 3



Gambar 3. data set setelah downsampling

3.2. Pre-processing

Tahap awal adalah pre-processing tahap ini diperlukan untuk memproses data agar lebih efektif [2] pada tahap pre-processing ada empat tahapan yaitu hilangkan angka di mana semua angka dari isi atribut tweet akan dihilangkan, lalu ada tahap lowercasting di mana mengubah semua huruf kapital menjadi huruf kecil, setelah itu ada tahap menghilangkan url di dalam data set masih terdapat url tweet dan itu tidak digunakan, setelah itu masuk ke tahap pembersihan dari tanda baca atau simbol karena tidak digunakan juga dalam proses bert dan klasifikasi

Tabel 2. Pre Processing

Isi Tweet sebelum Pre-Procssing	***Isi Tweet sesudah Pre-processing***
udaa 2 harian tidur gue ga nyenyak bgt, kaya gelisah gt. kenapa yaaa :([]	*udaa harian tidur gue ga nyenyak bgt kaya gelisah gt kenapa yaaa,1*
Ini gua kenapa ya...Udah mood berantakan trus tuh jantungnya kayak detaknya	ini gua kenapa yaudah mood berantakan trus tuh jantungnya detaknya kuat
kuat banget... gelisah banget rasanya.Ada sesuatu apa ya???.	banget gelisah banget kayak rasanyaada sesuatu apa ya
# saran lagu yang bikin tenang dong, gatau kenapa tiba-tiba muncul perasaan gusar/gelisah gini. Thanks!,1	*saran lagu yang bikin tenang dong gatau kenapa tibatiba muncul perasaan gusargelisah gini thanks,1*

3.3. Bert

Indobertweet-base-uncased merupakan sebuah model dari Bidirectional-Encoder Representasion (BERT) yang sudah dimodifikasi menyesuaikan penyematan sub kata dan beberapa tugas modifikasi [13] model ini dilatih menggunakan 26M Tweet dan memiliki 409M token kata [13]. Data yang dikumpulkan merupakan unggahan twitter yang dikumpulkan selama 1 tahun dari bulan Desember 2019 sampai Desember 2020 dengan 60 kata kunci dan 4 topik utama kesehatan, ekonomi, pemerintahan dan pendidikan [13].

Bert memakai Transformer untuk mempelajari hubungan kontekstual antar kata[11]. Berbeda dengan transformer biasa bert tidak bekerja sequential di mana membaca kalimat dari kiri ke kanan bert melihat keseluruhan hubungan antar kata dari kalimat tersebut[11].

Fine Tunning merupakan salah satu fitur /yang digunakan untuk melatih model BERT sesuai dengan permasalahan yang sedang dikerjakan[12]. Fine tunning dari indobertweet sudah pernah digunakan untuk sentimen, emosi dan klasifikasi ujaran kebencian. Fine tunning bert sendiri pernah digunakan untuk sentimen analisis dari drug review mendapatkan akurasi tertinggi yaitu 98.73% untuk data latih, 96.56% untuk data uji dan mendapatkan 90.67% dari data testing[14].

Pada penelitian ini fine tunning digunakan untuk tugas mengklasifikasikan mana kalimat mengandung kecemasan berlebih dan bukan.

Tabel 3. Parameter fine tunning

Parameter	***Nilai***
Num_epochs	*10*
Optimizer	*optimizer = torch.optim.AdamW(model_bert.parameters(), lr=2e-5)*
Criterion	torch.nn.CrossEntropyLoss()

Dalam proses *fine tunning* epoch diatur dalam jumlah 10 sementara untuk *optimizer* menggunakan AdamW dan *Learning Rate 2e-5* digunakan untuk *model bert* dan untuk *criterion* diisi untuk *loss* dari model *BERT.*

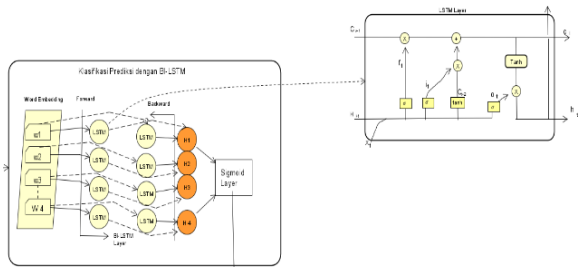
3.4. BI-LSTM

Lstm sendiri merupakan pengembangan dari metode *Recurent Neural Network* yang memiliki masalah dalam penanganan masalah mengingat

informasi panjang secara sequential[8] dan lstm dapat menangani masalah itu karna lstm menyimpan informasi jangka panjang[8].

Bidirectional – Long short term memory (BI-LSTM) bisa juga disebut lstm dua arah karna bi-lstm memiliki lapisan memori dari masa lalu variasi berarah ganda mempertimbangkan masukan dari kedua arah[6].

Bidirectional-Long Short Term Memory merupakan pengembangan dari pembelajaran mesin LSTM untuk melakukan sentimen pada kalimat yang mengandung stres atau depresi mendapatkan akurasi sebesar 0.90[5]. Bi-LSTM pernah digunakan untuk melakukan sentimen covid data didapatkan dari sosial media mendapatkan akurasi sebesar 95% untuk data latih dan 86% untuk data latih[9] dan Bi-lstm juga sudah pernah digunakan untuk mendeteksi unggahan berbentuk kalimat bermakna kecemasan berlebih mendapatkan akurasi sebesar 0.94 [8] sementara untuk *Long Short Term Memory mendapatkan akurasi sebesar 0.84%[8],* perbedaan Bi-lstm dan lstm adalah bi-lstm memiliki dua fitur yaitu *backward dan forward* di mana membaca kalimat dari awal ke akhir dan dari akhir ke awal.



Gambar 4. layer bi-lstm

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

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \quad (2)$$

$$C_t = tan\,h\,(W_c \cdot [h_{t-1}, X_t] + b_c) \quad (3)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot C_t \quad (4)$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o \quad (5)$$

$$h_t = O_t \cdot tan\,h\,(C_t) \quad (6)$$

$$h_t^{BiLSTM} = h_t^{forward} \oplus h_t^{backward} \quad (7)$$

Arsitektur Bi-lstm gerbang pertama adalah forget gate *di mana informasi akan digunakan atau dibuang pada dengan menggunakan persamaan (1). Lalu pada gerbang kedua ada* input gate *yang berfungsi untuk menyimpan informasi ke dalam* cell state *di dalam* input gate *ada dua lapisan yaitu* sigmoid *dan* tanh *untuk lapisan* sigmoid *berfungsi untuk memutuskan nilai mana yang akan disimpan di* cell state *seperti pada persamaan (2) sementara untuk* tanh *menambahkan nilai baru untuk* cell state *seperti pada persamaan (3). Keluaran dari untuk memperbaharui informasi dari* cell state. *Untuk persamaan (4) berfungsi untuk menghapus nilai di* forget gate *lalu ditambahkan* $i_t C_t$ *sebagai nilai baru dan akan masukan ke dalam cell state seperti* diuraikan di persamaan*(4). Selanjutnya adalah lapisan terakhir yaitu output gate di mana lapisan ini berfungsi* untuk menentukan *output cell state.* Yang pertama lapisan sigmoid akan menentukan bagian dari cell state yang akan dijadikan output seperti pada persamaan(5). Setelah itu *output* akan di masukan ke dalam lapisan tanh dan dikalikan dengan lapisan sigmoid agar menghasilkan *output* yang sudah diputuskan sebelumnya seperti diuraikan pada persamaan (6).

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
embedding_1 (Embedding)      (None, 64, 64)            105728

spatial_dropout1d_1 (Spati   (None, 64, 64)            0
alDropout1D)

bidirectional_1 (Bidirecti   (None, 512)               657408
onal)

dense_1 (Dense)              (None, 1)                 513

=================================================================
```

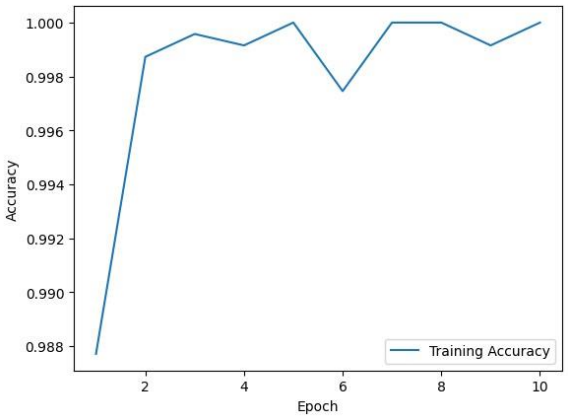
Gambar 5. Arsitektur Model Bi-LSTM

4. HASIL DAN PEMBAHASAN

Pada penelitian ini menggunakan data set dari *kaggle* di mana isi data set terdapat *url* tweet dan terdapat angka dan simbol oleh karena itu diperlukan *pre processing* untuk membersihkan isi dari unggahan tweet ada 4 tahapan di antaranya membersihkan url, membersihkan dari simbol, membersihkan dari angka, mengubah huruf menjadi kapital menjadi kecil dan menghilangkan spasi di awal dan akhir kalimat.

Untuk pembagian data set sendiri digunakan 70:30 untuk pelatihan fine tunning bert dan klasifikasi menggunakan *Bi-lstm.*

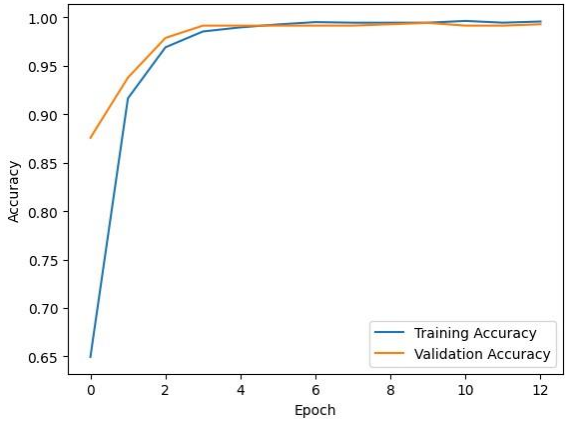
Proses *fine tunning bert* menggunakan model indobertweet mendapatkan akurasi sebesar 98.2% dari *training* data seperti pada gambar 5.



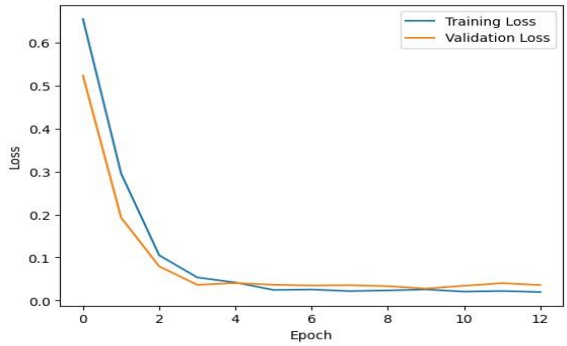
Gambar 6. Akurasi fine tunning bert

Sementara untuk klasifikasi bi-lstm sendiri sebelum masuk ke proses bi-lstm dimensi dari hasil *fine tunning bert* perlu melalui proses *Principal Component Analysist* untuk mengurangi dimensi dari

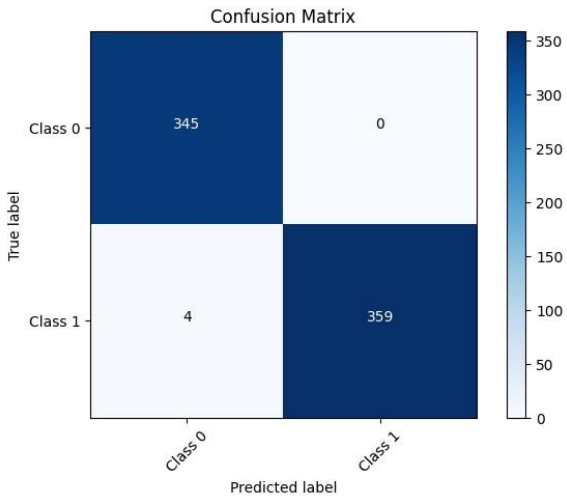
hasil *word embedding* bert yang berjumlah 768 dimensi diubah menjadi 64 dimensi dengan pembagian data latih dan data uji 70:30, 70 untuk data latih dan 30 untuk data uji mendapatkan akurasi sebesar 99.58% dari data latih berwarna biru sementara data uji 99.29% berwarna oren didapatkan dari 12 *epoch* dari epoch dari model yang sudah diatur 25 epoch sementara untuk loss sendiri dari data latih 0.0197 dan untuk data uji mendapatkan loss 0.0362.



Gambar 7. akurasi klasifikasi bi-lstm



Gambar 8. loss dari klasifikasi bi-lstm



Gambar 9. Confusion matrix

Data yang dipakai untuk melakukan uji confusion matrix merupakan data uji. Data uji sendiri di dapatkan dari pembagian jumlah data yaitu 70:30 pada tahapan awal data uji ini yang akan digunakan sebagai uji confusion matrix.

True Positive (TP): 345 - Ini adalah jumlah data uji yang di prediksi dengan benar sebagai positif (misalnya, di prediksi dengan benar sebagai Kelas 1).
False Positive (FP): 0 - Ini adalah jumlah data uji yang salah diprediksi sebagai positif (di prediksi sebagai Kelas 1 ketika kelas sebenarnya adalah 0).
False Negative (FN): 4 - Ini adalah jumlah data uji yang salah di prediksi sebagai negatif (di prediksi sebagai Kelas 0 ketika kelas sebenarnya adalah 1).
True Negative (TN): 359 - Ini adalah jumlah data uji yang di prediksi dengan benar sebagai negatif (misalnya, di prediksi dengan benar sebagai Kelas 0).
Akurasi = (TP + TN) / (TP + TN + FP + FN) = (345 + 359) / (345 + 0 + 4 + 359) ≈ 0.994
Presisi = TP / (TP + FP) = 345 / (345 + 0) = 1.0
Recall (Sensitivitas) = TP / (TP + FN) = 345 / (345 + 4) ≈ 0.988
Spesifisitas = TN / (TN + FP) = 359 / (359 + 0) = 1.0
Skor F1 = 2 * (Presisi * Recall) / (Presisi + Recall) = 2 * (1.0 * 0.988) / (1.0 + 0.988) ≈ 0.994

Penelitian yang sudah di kerjakan di mana deteksi unggahan berbentuk teks dari *Tweet* menggunakan metode *word embedding bert* untuk pembobotan kata dan menggunakan metode *Bidirectional-long short term memory* mendapatkan akurasi yang sangat bagus yaitu 99.44% didapatkan dari uji *confusion matrix* dengan loss 0.0278 dapat disimpulkan metode *word embedding bert* dapat meningkatkan akurasi dan mencegah model bi-lstm dari *overfitting* dikarenakan representasi numerik yang beragam dari proses *fine tunning word embeddings* dan juga dapat meningkatkan akurasi dari model *BI-LSTM*.

Untuk saran pengembangan penelitian ini bisa dikembangkan dengan menambahkan data set yang lebih beragam bentuknya tidak hanya teks tapi bisa juga video, gambar atau suara agar meningkatkan akurasi dari deteksi kecemasan berlebih dari sebuah postingan di media sosial dan juga diperlukan data set berjumlah besar agar model yang dilatih semakin bagus.

5. KESIMPULAN DAN SARAN

Kesimpulan dari penelitian yang sudah saya kerjakan adalah penerapan fitur *word embedding* dari BERT dan metode Bi-LSTM memberikan hasil akurasi yang tinggi dalam mendeteksi kalimat yang mengindikasikan kecemasan berlebih atau sebaliknya. Hal ini menunjukkan bahwa konversi teks menjadi vektor numerik melalui *word embedding*, khususnya dengan memanfaatkan fitur dari BERT dan pendekatan Bi-LSTM, dapat meningkatkan representasi semantik dari teks, dan pada gilirannya, meningkatkan akurasi dalam pengenalan pola kalimat yang bersifat emosional atau cemas. Hasil penelitian ini mendorong pemahaman lebih lanjut tentang

efektivitas teknik-teknik pengolahan teks, khususnya menggunakan *word embedding*, untuk meningkatkan representasi vektor teks. Kesimpulan ini menekankan bahwa penggunaan *word embedding*, terutama dengan mengintegrasikan fitur dari BERT dan menerapkan metode Bi-LSTM, dapat dianggap sebagai strategi efektif dalam meningkatkan akurasi representasi vektor teks. Oleh karena itu, temuan ini dapat memberikan kontribusi positif terhadap penelitian lanjutan di berbagai bidang yang melibatkan analisis teks dan pemahaman konten berbasis kalimat atau dokumen. Pada penelitian ini masih banyak yang harus dikembangkan seperti mencoba dengan data set dengan jumlah yang banyak dan lebih beragam. Bisa juga dikembangkan dengan menambahkan pengolahan gambar atau audio untuk mendeteksi seseorang terkena kecemasan berlebih atau tidak berdasarkan unggahan di sosial media atau bisa juga dikembangkan dengan menggunakan gabungan metode antara dari RNN dan CNN.

DAFTAR PUSTAKA

[1] Ghosal, S., & Jain, A. (2022). Depression and Suicide Risk Detection on Social Media using fastText Embedding and XGBoost Classifier. *Procedia Computer Science*, *218*, 1631–1639. https://doi.org/10.1016/j.procs.2023.01.141

[2] Singh, J., Wazid, M., Singh, D. P., & Pundir, S. (2022). An embedded LSTM based scheme for depression detection and analysis. *Procedia Computer Science*, *215*, 166–175. https://doi.org/10.1016/j.procs.2022.12.019

[3] Blanco, G., & Lourenço, A. (2022). Optimism and pessimism analysis using deep learning on COVID-19 related twitter conversations. *Information Processing and Management*, *59*(3). https://doi.org/10.1016/j.ipm.2022.102918

[4] Ghosh, T., Banna, M. H. al, Nahian, M. J. al, Uddin, M. N., Kaiser, M. S., & Mahmud, M. (2023). An attention-based hybrid architecture with explainability for depressive social media text detection in Bangla. *Expert Systems with Applications*, *213*. https://doi.org/10.1016/j.eswa.2022.119007

[5] Vandana, Marriwala, N., & Chaudhary, D. (2023). A hybrid model for depression detection using deep learning. *Measurement: Sensors*, *25*. https://doi.org/10.1016/j.measen.2022.100587

[6] Polignano, M., de Gemmis, M., Basile, P., & Semeraro, G. (2019). A comparison of Word-Embeddings in Emotion Detection from Text using BiLSTM, CNN and Self-Attention. *ACM UMAP 2019 Adjunct - Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, 63–68. https://doi.org/10.1145/3314183.3324983

[7] Ilham, F., & Maharani, W. (2022). Analyze Detection Depression In Social Media Twitter Using Bidirectional Encoder Representations from Transformers. *Journal of Information System Research (JOSH)*, *3*(4), 476–482. https://doi.org/10.47065/josh.v3i4.1885

[8] Setyo Nugroho, K., Akbar, I., & Nizar Suksmawati, A. (2021). *Seminar Nasional Hasil Riset Prefix-RTR DETEKSI DEPRESI DAN KECEMASAN PENGGUNA TWITTER MENGGUNAKAN BIDIRECTIONAL LSTM*.

[9] Arbane, M., Benlamri, R., Brik, Y., & Alahmar, A. D. (2023). Social media-based COVID-19 sentiment classification model using Bi-LSTM. *Expert Systems with Applications*, *212*. https://doi.org/10.1016/j.eswa.2022.118710

[10] Wu, J. L., He, Y., Yu, L. C., & Robert Lai, K. (2020). Identifying Emotion Labels from Psychiatric Social Texts Using a Bi-Directional LSTM-CNN Model. *IEEE Access*, *8*, 66638–66646. https://doi.org/10.1109/ACCESS.2020.2985228

[11] Tugas, J., Fakultas, A., Putra, H. K., Arif Bijaksana, M., & Romadhony, A. (n.d.). *Deteksi Penggunaan Kalimat Abusive Pada Teks Bahasa Indonesia Menggunakan Metode Indoin*.

[12] Choi, H., Kim, J., Joe, S., & Gwon, Y. (2020). Evaluation of BERT and Albert sentence embedding performance on downstream NLP tasks. *Proceedings - International Conference on Pattern Recognition*, 5482–5487. https://doi.org/10.1109/ICPR48806.2021.9412102

[13] Koto, F., Lau, J. H., & Baldwin, T. (2021). *IndoBERTweet: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization*. http://arxiv.org/abs/2109.04607

[14] Wang, X., Wang, X., & Zhang, S. (2023). Adverse Drug Reaction Detection From Social Media Based on Quantum Bi-LSTM With Attention. *IEEE Access*, *11*, 16194–16202. https://doi.org/10.1109/ACCESS.2022.3151900

[15] Zuo, L. Q., Sun, H. M., Mao, Q. C., Qi, R., & Jia, R. S. (2019). Natural Scene Text Recognition Based on Encoder-Decoder Framework. *IEEE Access*, *7*, 62616–62623. https://doi.org/10.1109/ACCESS.2019.2916616