

Analisis Sentimen Masyarakat terhadap Fenomena Childfree Menggunakan Metode Long Short Term Memory dan Bidirectional Encoder Representations from Transformers di Twitter

Zaemita Wahidatul Farida¹, Naim Rochmawati²

^{1,2} Jurusan Teknik Informatika/ Program Studi S1 Teknik Informatika, Universitas Negeri Surabaya

¹zaemita.19015@mhs.unesa.ac.id

²naimrochmawati@unesa.ac.id

Abstrak— Media sosial menjadi sarana bagi masyarakat untuk saling berinteraksi secara virtual. Penggunaan media sosial tidak terbatas pada berinteraksi saja, melainkan menjadi sarana pemanfaatan dalam melakukan penelitian, seperti halnya pada media sosial twitter. Twitter menjadi tempat yang ramai ketika terdapat isu terkini di negara ini bahkan dunia, beberapa isu terkini menjadi perhatian khusus oleh masyarakat, salah satunya tentang isu fenomena childfree, sehingga penelitian ini dilakukan dengan tujuan untuk mengetahui analisis sentimen terhadap fenomena menggunakan metode *Long Short Term Memory* dengan *Bidirectional Encoder Representations from Transformers* di twitter, ini menjadi kombinasi model yang dapat meningkatkan akurasi. Data yang digunakan yaitu berjumlah 3000 data dari twitter. Pada pengujian model dilakukan dengan pengujian beberapa parameter untuk mendapatkan model yang optimal diantaranya ukuran batch size 128, dropout 0.5, dense layer 32 dan lstm layer 64. Hyperparameter tersebut dilakukan pelatihan model yang menghasilkan performa terbaik dengan akurasi sebesar 0.9585, f1-score 0.9589, loss 0.1001, kemudian model dapat memprediksi tweets childfree dan menghasilkan precision sebesar 0.7839, recall 0.77, dan f1-score 0.7697.

Kata Kunci— *Childfree, Twitter, Sentimen, Long Short Term Memory, Bidirectional Encoder Representations from Transformers*

I. PENDAHULUAN

Perkembangan teknologi membawa kemudahan di berbagai bidang dalam kehidupan manusia, terutama dalam bidang informasi dan komunikasi. Media sosial merupakan sarana bagi pengguna untuk saling berinteraksi dengan informasi yang ada secara virtual. Media sosial adalah medium di internet yang memungkinkan pengguna merepresentasikan dirinya maupun berinteraksi, bekerja sama, berbagi, berkomunikasi dengan pengguna lain membentuk ikatan sosial secara virtual [1]. Setiap jenis media sosial memiliki focus atau keunggulan masing-masing. Media sosial bukan hal yang asing di kehidupan sehari-hari, melalui media sosial seseorang dapat secara bebas mengekspresikan ide pikiran gagasan, kondisi perasaan, melaporkan kejadian yang terjadi, hingga mengirim kritik dan saran. Kritikan dicuatkan ke media sosial mengundang banyak reaksi dari sekalgan orang pengguna media sosial, salah satunya pada platform media sosial twitter, twitter merupakan media sosial yang digunakan untuk mengekspresikan opini dan perasaan dengan secara singkat

padat dan jelas karena pada twitter thread dibatasi 140 karakter saja, dengan batasan karakter ini pengguna dapat lebih ekspresif dibanding dengan menggunakan media sosial lainnya. Twitter menjadi tempat yang ramai ketika terdapat isu terkini di negara ini bahkan dunia, beberapa isu terkini menjadi perhatian khusus oleh masyarakat, salah satunya tentang isu fenomena childfree.

Childfree secara bahasa berarti bebas anak. Childfree merupakan keputusan untuk tidak melahirkan atau mempunyai keturunan yang disepakati setelah menikah [2]. Keputusan untuk childfree di negara Indonesia memang tidak ada larangan, tetapi berdasarkan keputusan ini akan berdampak pada sumber daya manusia di negara Indonesia khususnya yaitu kurangnya usia produktif di masa mendatang, dengan berkurangnya usia produktif ini akan berdampak pada pertumbuhan ekonomi di Indonesia.

Fungsi twitter menjadi lebih luas diakibatkan popularitas penggunaan yang cukup tinggi sehingga fungsi twitter yang awalnya sebagai platform virtual untuk mengekspresikan opini, tulisan atau tanya jawab dan berbagi informasi kini menjadi sarana yang digunakan di berbagai aspek salah satunya sebagai media penelitian berupa text mining, sentiment analysis, dan artificial intelligence [3]. Karakteristik data yang dimiliki twitter dengan jumlah yang banyak dan topik yang bervariasi membuat twitter memiliki potensi untuk dilakukan analisis sentimen.

Analisis sentimen adalah proses menganalisis teks digital, memahami, mengekstrak data opini untuk mendapatkan sebuah sentimen yang terkandung dalam sebuah opini dan menentukan apakah nada emosional pesan tersebut positif, negative, atau netral [4], sehingga dengan analisis sentimen fenomena childfree dapat mengetahui respon masyarakat terhadap keputusan untuk childfree melalui komentar pada postingan di media sosial. Sentimen analisis termasuk dari salah satu model klasifikasi data menggunakan pendekatan supervised learning di dalam machine learning. Machine learning memiliki 2 tipe teknik yaitu supervised learning dan unsupervised learning [5]. Analisis sentimen dapat diklasifikasikan menggunakan model machine learning atau deep learning. Deep learning dapat menganalisa dengan kompleks pada data yang sederhana dengan cara memproses banyak layer informasi nonlinear untuk melakukan proses ekstraksi fitur, pengenalan pola, dan klasifikasi (Yu &Deng, 2014). Dengan kinerja yang sangat baik

penggunaan deep learning dapat menanggapi klasifikasi sentimen pada teks, salah satu metode terbaik dari deep learning ini yaitu Long Short Term Memory (LSTM). LSTM memiliki struktur yang disebut gerbang (gates) berguna untuk mengatur aliran informasi. LSTM menghasilkan akurasi yang cukup baik untuk prediksi sentimen dibandingkan dengan beberapa model lain, tetapi hal itu tidak dapat dipungkiri bahwa mesin tidak mudah dalam memahami bahasa manusia, sehingga pengembangan model diperlukan agar hasil klasifikasi tidak mengandung kesalahan penafsiran pada objek sentimen Childfree, yaitu dengan menggunakan metode penghasil vektor kata Bidirectional Encoder Representations from Transformers (BERT). BERT bekerja secara sekuensial dan akan sesuai dengan deep learning yang merupakan word embedding, yang akan menghasilkan fitur berdimensi rendah tetapi memiliki informasi yang padat. Pada penelitian yang akan dilakukan menggunakan pertained BERT yang dapat meningkatkan akurasi karena telah dilatih oleh Wikipedia.

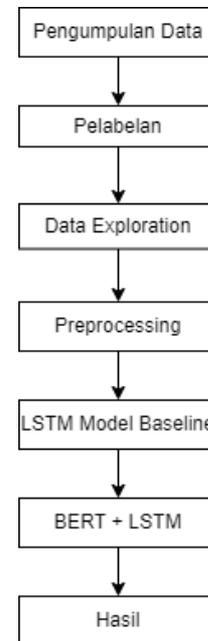
Contoh penelitian oleh Laina Farsiah dengan judul “Analisis Sentimen Menggunakan Arsitektur Long Short Term Memory (LSTM) terhadap Fenomena Citayem Fashion Week” [6]. Penelitian dilakukan jumlah training data sejumlah 80% dari total dataset 6.043 dan jumlah testing 20% dari total dataset 1.511. Model dilatih menggunakan beberapa konfigurasi hyperparameter tidak berubah, yaitu adam sebagai optimizer, jumlah epoch sebanyak 50, presentase dropout sebesar 0.2 dan konfigurasi hyperparameter diperbarui yaitu batch size sebanyak 128, learning rate sebesar 0.001 dan embed dimension sebesar 64 pada struktur Long Short Term Memory dengan hasil akurasi terbaik yaitu 88%. Kemudian penelitian yang dilakukan oleh David, et al dengan judul “Peningkatan Akurasi Klasifikasi Ulasan Makanan Dengan Bidirectional LSTM dan BERT Embedding”[7]. Dibuktikan bahwa kombinasi metode Bi-LSTM dan BERT Embedding dapat meningkatkan hasil akurasi dalam mengklasifikasi sentiment, dikarenakan model dapat lebih mengerti pola kalimat sehingga dapat mengurangi kesalahan klasifikasi dalam ulasan makanan Amazon.

Berdasarkan penelitian yang telah dilakukan sebelumnya, hasil penelitian menunjukkan bahwa penggunaan metode Long Short Term Memory (LSTM) dan Bidirectional Encoder Representations from Transformer (BERT) memiliki hasil akurasi yang tinggi dibandingkan dengan metode yang lain, dengan ini penulis tertarik untuk melakukan analisa kombinasi Machine Learning yaitu LSTM dan BERT untuk mengetahui tingkat akurasi yang dihasilkan apabila menggunakan metode Long Short Term Memory dan BERT Embedding di media sosial Twitter dengan judul “Analisis Sentimen Masyarakat terhadap Fenomena Childfree Menggunakan Metode Long Short Term Memory dan Bidirectional Encoder Representations from Transformers di Twitter”.

II. METODE PENELITIAN

Metode penelitian adalah suatu cara ilmiah untuk mendapatkan data dengan tujuan tertentu. Cara ilmiah berarti

kegiatan penelitian yang didasarkan pada ciri-ciri keilmuan yaitu rasional, empiris dan sistematis. [8]. Berdasarkan pengertian tersebut, dapat disimpulkan bahwa setiap permasalahan dapat diselesaikan dengan cara ilmiah. Pemilihan metode juga akan mempengaruhi hasil penelitian, sehingga dibutuhkan pemilihan metode yang tepat di setiap penelitiannya.



Gbr. 1. Alur Penelitian

Dalam melakukan klasifikasi analisis sentimen terdapat beberapa tahapan atau alur penelitian yang akan dilakukan mulai dari pengumpulan data hingga pemodelan dan evaluasi model seperti yang telah diuraikan pada Gbr. 1. Berikut penjelasan untuk setiap tahapan antara lain:

A. Pengumpulan Data

Pengumpulan data dilakukan dengan proses scraping data pada media sosial twitter dengan menggunakan Tools Google Colaboratory dan Bahasa pemrograman python. Pengambilan data menggunakan *library snsrape*, dimana snsrape menyediakan API sederhana (API GraphQL Twitter) untuk mengakses data Twitter secara langsung. Data yang dihasilkan berisi Komentar Tweets dengan kata kunci pencarian #gitasav , #childfree.

B. Pelabelan

Pelabelan dilakukan setelah data diperoleh dengan tujuan untuk mengklasifikasi data menjadi kategori sentimen yang berbeda, yaitu positif, negative, dan netral. Pelabelan dilakukan untuk memahami dan menggambarkan pandangan atau perasaan yang terkandung dalam sentiment childfree.

Pelabelan dilakukan secara manual sehingga dapat diperiksa dan direvisi berdasarkan tambahan pengetahuan yang dimiliki oleh pelabel sehingga ini akan memungkinkan untuk memperbaiki serta meningkatkan kualitas label. Berikut disajikan contoh pelabelan pada Gbr.2.

1	Pilih pada kolom Sentimen di bawah apakah pernyataan tweet yang bersangkutan menurut Anda negatif (tidak setuju), netral, atau positif (setuju) terhadap kasus child free pada gitasav belakangan ini. Jika sudah yakin dengan jawaban tersebut, silahkan ganti pada kolom Yakin.	3000
2	Komentar	Sentimen
3	asanlita Km org nangkepnya anak 3 pembahasan childfree gitasav kali ya jadi dianggap awet muda padahal ya emang masih muda terlepas di luar sana mgkn ada yg seumuran iris tp keliatan lebih tua Shes pretty and her aura is chefs kiss	Negatif
4	buat apa pake walau tidak childfree tho i dont really know a lot about that gitasav girllie but if the whole country keep dragging her for the option she chose for her own body then i suggest yall better move on and focus on your own lives	Netral
5	lya ih Naon sih Awet muda walau tidak childfree Susah emg tetangga Gitasav nih Pola pikirnya sama aja	Negatif
6	andrihw_cicitraa tang__kira Ya itu alamiahnya ada background factors yang bikin mereka begitu Jadi ga aneh punya dampak sangat signifikan Respon atas reply yang bilang bahaya ya ternyata childfree ya kalo childfree nya minoritas	Netral
7	asanlita Si mbak ini awalnya cuma di sandiing sama gitasav sebagai perbandingan dulu pas berita childfree seliweran mangkannya masih paham point awalnya selebihnya yg ini sudah oot semua	Netral
8	Masih ga paham kenapa orang-orang segitunya sm gitasav padahal dia ga ada nyindir di tulisan ttg childfree nya	Netral

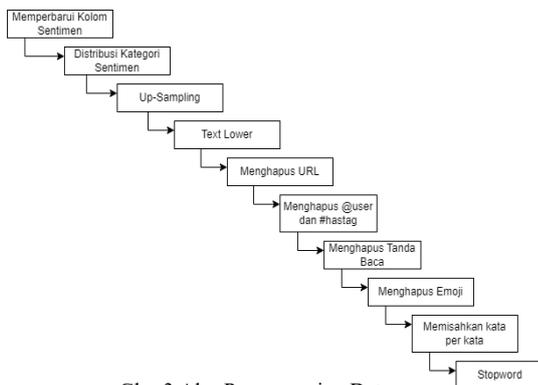
Gbr. 2 Contoh Pelabelan

C. Data Exploration

Proses Eksplorasi data yang dilakukan untuk menganalisis teks untuk memahami karakteristik, mencari pola dan mengidentifikasi wawasan yang diperlukan, seperti menghitung distribusi panjang kalimat, menghitung percentile untuk menentukan *seq_len*, menghitung kemunculan kata terbanyak, melakukan visualisasi data, dan menjalajahi hubungan antara kata-kata atau urutan kata dalam teks.

D. Preprocessing Data

Preprocessing dilakukan untuk menyiapkan data untuk pemodelan seperti membersihkan data dari unsur yang tidak berguna dan agar data lebih mudah untuk dipahami oleh sistem pengolahan text mining.



Gbr. 3 Alur Preprocessing Data

Adapun penjelasan dari beberapa tahap dalam proses Preprocessing data sebagai berikut.

1) *Memperbarui Kategori Sentimen*: memperbarui kategori sentimen dengan mengubah semua nilai dalam sentimen menjadi huruf kecil dan menghapus spasi ekstra di awal dan akhir dalam sentimen.

2) *Distribusi Kategori Sentimen*: proses enkripsi nilai sentimen dialihkan menjadi nilai categorical menggunakan pengkodean label.

3) *Up-Sampling*: mencapai keseimbangan jumlah sampel antara kelas minoritas (positif) dan kelas minoritas (negatif dan netral), kelas negative dan netral akan ditingkatkan jumlah sampelnya sebanding dengan jumlah sampel kelas positif.

4) *Text Lower*: mengubah seluruh teks tweet menjadi huruf kecil

5) *Cleaning*: menghapus elemen yang tidak digunakan seperti menghapus URL, menghapus @username, karakter '#', tanda baca, dan emoji

6) *Tokenizing*: memisahkan teks menjadi kata per kata atau token individu.

7) *Stopword*: menghapus kata yang tidak diperlukan atau tidak bermakna secara relevan.

E. Pemodelan

Pada tahap pemodelan, data yang telah siap dan dilakukan preprocessing akan dimasukkan ke dalam model. Model yang digunakan pada penelitian ini adalah gabungan model *LSTM* dengan model *pre-trained BERT*, dengan melakukan kombinasi kedua model dapat memungkinkan model untuk memahami makna kata dan konteks yang merupakan kelebihan *BERT* serta kemampuan *LSTM* dalam menangani urutan panjang dependensi jarak jauh, berdasarkan keunggulan masing-masing model tersebut sehingga dapat dilakukan kombinasi yang baik dan mendapatkan hasil yang memuaskan dibanding model lain.

Tahap pemodelan dibagi menjadi beberapa tahap dengan penjelasan sebagai berikut.

1) Set model BERT *pre-trained*

Dilakukan instalasi BERT *pre-trained*. Model BERT yang digunakan yaitu *indo-bert-base-indonesian-522M* yang telah dilatih dengan kumpulan data Bahasa Indonesia, Wikipedia.

2) Tokenizer

Proses tokenizer menggunakan library *hugging face transformers* sehingga akan menghasilkan embedding dan dimuat secara otomatis menggunakan *pre-trained tokenizer* kemudian disimpan dalam 'bert_model'.

3) Create Dataset

Tahap selanjutnya adalah *create dataset* menggunakan Tensorflow. Dataset berisi tiga elemen untuk setiap sampel; input IDs, mask, dan label sentiment dalam bentuk one-hot-encoding. Setiap kata akan diwakili oleh vector biner dimana hanya satu elemen dari vektor akan

bernilai 1 mewakili kata itu sendiri, dan elemen lainnya adalah 0. Contoh {"saya","setuju","dengan","childfree"}, maka representasi "setuju" dalam bentuk one-hot encoding menjadi [0,1,0,0].

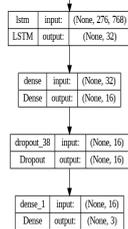
4) Split Dataset

Pembagian dataset dilakukan dengan membagi dataset menjadi subset pelatihan (*training*) dan subset validasi (*validation*) dengan proporsi 80% : 20%. 80% data training dan 20% data validation. Proporsi pembagian dataset yang digunakan pada penelitian ini ditentukan berdasarkan hasil terbaik dari pengujian beberapa proporsi yang berbeda.

5) Create Model BERT with LSTM



bert	input:	(None, 276)
TFBertMainLayer	output:	TFBaseModelOutputWithPoolingAndCrossAttentiones[enc_hidden_state=(None, 276, 768), pooler_output=(None, 768), past_key_values=None, hidden_states=None, attentions=None, cross_attentions=None]



Gbr. 4 Arsitektur BERT & LSTM

Pada tahap ini penggabungan model BERT dengan LSTM dilakukan. Arsitektur model dibuat dengan model BERT sebagai lapisan awal dan diteruskan ke lapisan LSTM seperti pada gambar 4, pada model BERT tersebut dilakukan pendefinisian input layer yang digunakan untuk merepresentasikan *token* dan *attention mask* dalam *TensorFlowKeras*, dilanjutkan dengan proses embedding dengan mengambil token dan daftar mask, selanjutnya model BERT dijalankan untuk menghasilkan representasi embedding. Selanjutnya dilakukan pemodelan LSTM tunggal yang telah dikurangi kompleksitasnya. Model LSTM yang telah dilakukan *complexity* ini akan menerima embedding teks dari model BERT yang telah dijelaskan sebelumnya sebagai input dan akan menghasilkan keluaran dalam bentuk probabilitas distribusi kelas yang dihasilkan oleh fungsi softmax.

6) Kompilasi Model

Proses kompilasi model dimulai dengan perhitungan *f1-score* menggunakan fungsi (*f1_score(y_true,y_pred)*) yang dibuat untuk menghitung nilai *f1-score* dengan menggunakan *Keras backend(K)*. Kompilasi model selanjutnya yaitu melakukan pengaturan *optimizer*, pengaturan *optimizer* yang dimaksud yaitu mengatur *loss*

function dan *metrics* yang akan digunakan untuk pelatihan model. *Optimizer* yang digunakan yaitu optimizer Adam dengan *learning rate* 0.0005, penggunaan *CategoricalAccuracy* sebagai fungsi kerugian untuk meminimalkan kesalahan klasifikasi antar kelas, dilanjut mengompilasi model dengan menggabungkan *optimizer*, *loss function*, dan *metrics*.

III. HASIL DAN PEMBAHASAN

A. Data

Data yang dikumpulkan yaitu data komentar tweets dengan teknik pengumpulan data (*scrapping*) menggunakan *library snsrape* yang diambil dari Twitter. Total data yang dikumpulkan disajikan pada tabel I berikut.

TABEL I
JUMLAH DATA HASIL SCRAPING

Sumber Data	Jumlah	Total
Twitter	7617	7617

Data yang dikumpulkan sejumlah 7617 tweets merupakan data tweets mentah yang belum terstruktur, dan terdapat noises. Setelah data terkumpul selanjutnya dilakukan preprocessing dan diberi label, hasil label disajikan pada tabel 2. Proses pelabelan pada penelitian ini dilakukan secara manual menggunakan kategori sentiment positif, negatif, dan netral. Tweet yang memuat tulisan atau komentar positif terhadap *childfree* dan/atau mendukung keputusan *childfree* dikategorikan dengan label positif, sebaliknya apabila tweet berisi umpatan, memberikan komentar tidak setuju dengan *childfree*, maka tweet tersebut dikategorikan dengan label sentiment negatif, dan tweet yang tidak memberikan pendapat setuju atau tidak setuju terhadap *childfree*, maka dikategorikan dengan label sentiment netral.

TABEL II
ISI HASIL SCRAPING

Komentar	Sentimen
#gitasav was right lebih baik gue <i>childfree</i>	positif
I mean, se-open minded nya gue masih mau ko punya anak	negatif
Childfree itu bukan budaya kita, budaya kita itu banyak anak banyak rezeki	negatif
I still dont get how and why do people really care about other people for the	netral

agar tidak dapat diubah, bobot LSTM tidak di-update selama proses pelatihan, sehingga pembuatan model menjadi lebih sederhana. Setelah pemodelan dilakukan, selanjutnya dilakukan proses pelatihan dengan pengaturan parameter sebagai berikut.

- *Batch size* : 128
- *Lstm layer* : 64
- *Dense layer* : 32
- *Dropout* : 0,5
- *Learning rate* : Adam 0,0005

Pada tahap pemodelan ini menggunakan *batch size*: 128, penggunaan jumlah ini dibuktikan oleh hasil hasil terbaik dari pengujian penggunaan jumlah batch size yaitu antara 32, 64, dan 128. Penggunaan jumlah batch size ini berdasarkan pertimbangan beberapa hal seperti ukuran dataset pada penelitian, dan kompleksitas model yang digunakan. Pada penelitian ini model LSTM dan BERT termasuk kombinasi yang cukup kompleks, sehingga membutuhkan batch size yang cukup besar untuk memaksimalkan efisiensi pelatihan dan menghindari *overfitting*.

Berikut ini pada tabel 7 merupakan hasil pengujian berdasarkan jumlah split dataset, pengujian ini dilakukan untuk mengetahui hasil terbaik dengan jumlah split dataset tertentu, dengan jumlah proporsi yaitu 80:20, 75:25, 70:30, 65:35, 60:40.

TABEL VII
HASIL PENGUJIAN SPLIT DATASET

Split Data	Accuracy	Val_Accuracy	F-1 Score	Val_F-1 Score	Loss	Val_Loss
80:20	0.8862	0.8964	0.8838	0.8989	0.3055	0.2474
75:25	0.8690	0.8734	0.8634	0.8714	0.3523	0.3252
70:30	0.8481	0.7879	0.8438	0.7846	0.3877	0.4930
65:35	0.9245	0.8750	0.9233	0.8756	0.2102	0.3319
60:40	0.8438	0.7760	0.8437	0.7760	0.4000	0.5807

Pengujian berdasarkan jumlah proporsi pembagian data training dan validation ini digunakan untuk membuktikan bahwa pembagian dataset dengan jumlah proporsi 80:20 yang meliputi 80% data pelatihan dan 20% data validasi pada penelitian ini merupakan pembagian dataset terbaik. Berdasarkan pengujian diatas dapat disimpulkan bahwa pembagian dataset 80:20 mendapatkan hasil terbaik, sehingga pembagian dataset dengan proporsi 80:20 digunakan pada penelitian ini.

Setelah diketahui pembagian dataset terbaik, selanjutnya dilakukan pengujian berdasarkan kombinasi parameter meliputi ukuran batch size, lstm layer, dropout, dense layer,

sehingga dapat diketahui hasil hyperparameter tuning terbaik dalam pengujian model.

1. Pengujian 1

Pengujian 1 yaitu menggunakan jumlah batch size 64, lstm layer 32, dense layer 16 dan dropout 0.3, dengan hasil disajikan pada tabel 8 berikut.

TABEL VIII
HASIL PENGUJIAN 1

Accuracy	0.8810
Val_Accuracy	0.9583
F-1 Score	0.8738
Val_F1-Score	0.9597
Loss	0.5378
Val_Loss	0.3953

2. Pengujian 2

Pengujian 2 yaitu menggunakan jumlah batch size 128, lstm layer 64, dense layer 32 dan dropout 0.5, dengan hasil disajikan pada tabel 9 berikut.

TABEL IX
HASIL PENGUJIAN 2

Accuracy	0.9340
Val_Accuracy	0.9585
F-1 Score	0.9335
Val_F1-Score	0.9589
Loss	0.2051
Val_Loss	0.1001

3. Pengujian 3

Pengujian 3 yaitu menggunakan jumlah batch size 128, lstm layer 64, dense layer 64, 32 dan dropout 0.5, dengan hasil disajikan pada tabel 10 berikut.

TABEL X
HASIL PENGUJIAN 3

Accuracy	0.9196
Val_Accuracy	0.9456
F-1 Score	0.9193
Val_F1-Score	0.9544
Loss	0.2443
Val_Loss	0.1571

Pengujian 2 dengan pengaturan parameter yaitu jumlah batch size 128, lstm layer 64, dense layer 32 dan dropout 0.5 mendapat hasil terbaik dibandingkan dengan pengujian 1, dan 3, dengan ini berdasarkan pengujian yang telah dilakukan dapat diketahui bahwa pemilihan penggunaan parameter berpengaruh terhadap akurasi dari pengujian suatu model, seperti contoh penggunaan batch size, batch size merupakan factor penting dalam pengelolaan proses pelatihan model.

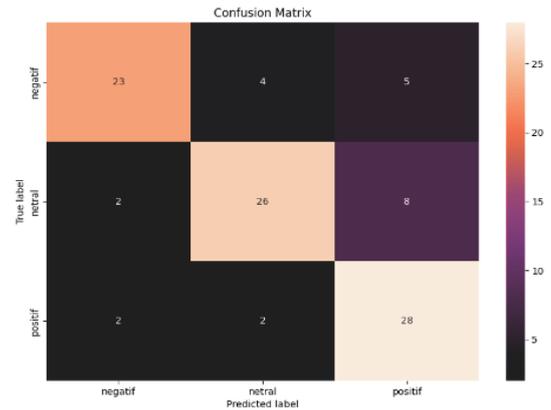
Berikut dilakukan pengujian penggunaan parameter batch size untuk mengetahui berapa jumlah batch size yang terbaik untuk pelatihan model pada penelitian ini. Hasil dari pengujian penggunaan batch size yaitu dengan jumlah batch size 32, 64, dan 128, disajikan pada tabel 11 berikut.

TABEL XI
HASIL PENGUJIAN BATCH SIZE

Batch Size	Accuracy	Val_Accuracy	F-1 Score	Val_F-1 Score	Loss	Val_Loss
128	0.8862	0.8964	0.8838	0.8989	0.3055	0.2474
64	0.8678	0.8872	0.8618	0.8827	0.3629	0.2868
32	0.8333	0.8090	0.8273	0.8113	0.4388	0.4545

Dari pengujian batch size diatas, diketahui hasil terbaik dari penggunaan batch size yaitu pada jumlah batch size 128, sehingga jumlah batch size 128 digunakan pada penelitian yang dilakukan seperti pada pengujian 2 yang mendapat hasil terbaik dibandingkan dengan pengujian 1, dan 3.

Setelah dilakukan beberapa pengujian dan diketahui hasil terbaik pada pengujian 2, selanjutnya dilakukan evaluasi pada model. Pada tahap ini model akan diuji untuk memprediksi kelas sentimen dengan menggunakan data uji berjumlah 100 tweets, data ini merupakan data yang berbeda dari data pada pelatihan yang digunakan sebelumnya yaitu data yang belum pernah dilihat oleh mesin sebelumnya. Tujuan dilakukan evaluasi yaitu untuk melihat seberapa tinggi tingkat kepercayaan model dalam memprediksi kelas sentimen. Hasil evaluasi confusion matrix disajikan pada gambar 6 dengan hasil perhitungan evaluasi pada tabel 12 berikut.



Gbr. 6 Confusion Matrix Predict

Dari gambar diatas dapat dijelaskan bahwa nilai yang memprediksi sentimen positif dan hasil sebenarnya adalah positif yaitu sebanyak 28 data, prediksi negatif dengan hasil sebenarnya positif adalah 2 data, prediksi netral dengan hasil sebenarnya positif yaitu sebanyak 2 data, kemudian pada prediksi negative dengan hasil sebenarnya negative adalah 23 data, prediksi positif dengan hasil sebenarnya adalah negative yaitu 5 data, prediksi netral tetapi hasil sebenarnya adalah negative adalah sebanyak 4 data, selanjutnya pada prediksi netral dengan hasil sebenarnya adalah netral adalah 26 data, prediksi positif dengan hasil sebenarnya adalah netral sebanyak 8 data, dan prediksi negatif tetapi hasil sebenarnya adalah netral adalah 2 data. Berdasarkan nilai prediksi sentiment diatas dapat disimpulkan hasil perhitungan evaluasi pada model yang disajikan pada tabel 12 berikut.

TABEL XII
HASIL EVALUASI MATRIX

Evaluasi	Hasil
Pressicion	0.7839
Recall	0.77
F1-Score	0.7697

Pada tahap selanjutnya dilakukan prediksi label untuk mengetahui seberapa tinggi nilai kepercayaan pada label yang telah diberikan pada komentar tweet secara manual sebelumnya. Berikut hasil prediksi label disajikan pada tabel 13 berikut.

TABEL XIII
HASIL PREDIKSI LABEL

Tweet	True label	Predicted label	Confidence
gitasav emang benar sih lebih baik childfree saja kalau ga bias didik anak dan anaknya jadi gini	positif	positif	99.86186

untuk kasus tertentu childfree yang dianut gitasav ini ada benarnya	positif	netral	61.35224
pemikiran gitasav tentang childfree tidak cocok untuk saya yang mau punya anak banyak yang suka mengurus anak yang suka main sama anak	negatif	positif	89.88631
tuhan yang gamau punya anak gitasav doing akum ah mau tuhan	negatif	negatif	58.92666

IV KESIMPULAN

Analisis sentimen komentar masyarakat tentang topik childfree di Twitter pada penelitian ini yaitu berjumlah 3000 data dengan jumlah kelas positif 1110, negatif 1070, dan netral 820. Data tersebut diperoleh melalui proses scraping twitter dan telah dilakukan proses pengelolaan data, hingga pengimplementasian data pada model yaitu *Long Short Term Memory (LSTM)* dan *Bidirectional Encoder Representations from Transformers (BERT)*. Kombinasi metode dilakukan pada penelitian ini dengan tujuan untuk mendapat hasil terbaik pada analisis sentimen yang dilakukan, pada hasil pelatihan yang dilakukan mendapatkan hasil nilai akurasi sebesar 0.9585 nilai f-1 score 0.9589 dengan nilai kerugian sebesar 0.1001. Selanjutnya pada hasil prediksi twwet childfree yaitu menghasilkan nilai precision sebesar 0.7839, recall 0.77, dan f1-score 0.7697. Berdasarkan hasil tersebut dapat disimpulkan bahwa metode LSTM dan BERT dapat dikombinasikan untuk dilakukan implementasi pada analisis sentimen, dalam hal ini analisis sentimen terhadap fenomena childfree.

IV SARAN

Berdasarkan penelitian yang sudah dilakukan, saran yang dapat diberikan untuk penelitian selanjutnya adalah:

1. Penambahan jumlah dataset dapat dilakukan untuk mengetahui pengaruh dari jumlah dataset terhadap hasil akurasi
2. Pengaturan parameter yang beragam untuk mengetahui pengaruh dari parameter yang digunakan pada pelatihan model
3. Dapat dibuat sistem analisis sentiment yang berguna untuk memprediksi komentar sebagai survey untuk berbagai fenomena yang ada di Indonesia.

UCAPAN TERIMA KASIH

Puji syukur kepada Allah SWT yang telah memberikan kesehatan, rezeki berupa ilmu, dan kemudahan lainnya sehingga penulis dapat menyusun artikel ilmiah ini dengan baik. Terima kasih kepada kedua orang tua beserta keluarga yang selalu memberikan dukungan dan doa secara penuh kepada penulis, serta terima kepada sahabat penulis yang menemani, memberi masukan atau bantuan selama melakukan penelitian. Tidak akan terlupa terima kasih kepada diri sendiri yang telah kuat, bertahan dan berjuang untuk menyelesaikan artikel ilmiah ini dengan baik.

REFERENSI

- [1] Nasrullah, R. (2015). Media sosial: perspektif komunikasi, budaya, dan sosioteknologi. Bandung: Simbiosis Rekatama Media.
- [2] Makhlad, M. (2022). Analisis Childfree Dalam Website Tafsir Di Indonesia (Studi Atas TafsirQuran. Id, Tanwir. Id, Dan Mubadalah. Id) (Bachelor's thesis, FU).
- [3] Hartanto, H. (2017). Text Mining dan Sentimen Analisis Twitter pada Gerakan LGBT. Intuisi: Jurnal Psikologi Ilmiah, 9(1), 18-25.
- [4] Masruroh, S. U., & Wardhani, L. K. Analisis kinerja algoritma naive bayes dan k-nearest neighbor pada sentimen analisis dengan pendekatan lexicon di media twitter (Bachelor's thesis, Fakultas Sains dan Teknologi Universitas Islam Negeri Syarif Hidayatullah Jakarta).
- [5] Brownlee, J. (2016). Master Machine Learning Algorithms: discover how they work and implement them from scratch. Jason Brownlee.
- [6] Farsiah, L., Misbullah, A., & Husaini, H. (2022). Analisis Sentimen Menggunakan Arsitektur Long Short-Term Memory (Lstm) Terhadap Fenomena Citayam Fashion Week. Cyberspace: Jurnal Pendidikan Teknologi Informasi, 6(2), 86-94.
- [7] Pasaribu, D. J. M., Kusriani, K., & Sudarmawan, S. (2020). Peningkatan Akurasi Klasifikasi Sentimen Ulasan Makanan Amazon dengan Bidirectional LSTM dan Bert Embedding. Inspiration: Jurnal Teknologi Informasi dan Komunikasi, 10(1), 9-20.
- [8] Darmadi, Hamid. 2013, Metode Penelitian Pendidikan dan Sosial, Alfabeta, Bandung.