

Analisis Sentimen Sebagai Deteksi Sarkasme di Media Sosial X Berbahasa Indonesia Menggunakan Transformer dengan Model XLNet

Haqul Amal Jiddan¹, Ricky Eka Putra²

^{1,2}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Negeri Surabaya

¹haqul.21028@mhs.unesa.ac.id

²rickyeka@unesa.ac.id

Abstrak— Media sosial, khususnya X, sering memuat ujaran sarkastik yang sulit dikenali tanpa pemahaman konteks. Penelitian ini bertujuan mendeteksi sarkasme pada teks berbahasa Indonesia di X menggunakan model berbasis transformer XLNet. Data diambil dari Hugging Face dan melalui preprocessing untuk memastikan konsistensi. Kemudian, data dibagi ulang menjadi tiga skenario yaitu 70:15:15, 60:20:20, dan 80:10:10. Model dilatih (fine-tuning) dengan variasi learning rate, batch size, dropout, sequence length, dan epoch. Pemilihan kandidat mempertimbangkan macro-F1 dan recall pada label sarkasme. Hasil menunjukkan pada skenario 70:15:15, XLNet percobaan ke-17 menjadi model terbaik pada penelitian ini, meraih Accuracy 0.81 dan Macro-F1 0.81 dan pada classification report pada label non-sarkasme memperoleh Precision 0.82, Recall 0.78, dan F1-Score 0.80. Lalu, pada label sarkasme memperoleh Precision 0.79, Recall 0.83, F1-Score 0.81). Temuan ini diharapkan menjadi referensi pengembangan sistem analisis opini atau ujaran sarkastik pada platform media sosial di Indonesia.

Kata Kunci— Analisis sentimen, Sarkasme, XLNet, X, Transformer

I. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi menjadikan media sosial sebagai salah satu sarana utama dalam menyampaikan pendapat dan informasi. X, sebagai salah satu platform media sosial terpopuler di Indonesia, menjadi wadah bagi masyarakat untuk mengekspresikan opini mereka tentang berbagai topik, baik yang bersifat pribadi maupun publik. Salah satu bentuk ekspresi yang sering digunakan dalam interaksi di media sosial adalah sarkasme. Sarkasme kerap digunakan untuk menyampaikan humor, kritik, atau ketidakpuasan secara tidak langsung. Namun, karena sifatnya yang ambigu, kontekstual, dan sering kali tersamar, sarkasme sulit dipahami dan dapat menyebabkan kesalahpahaman [1], [2].

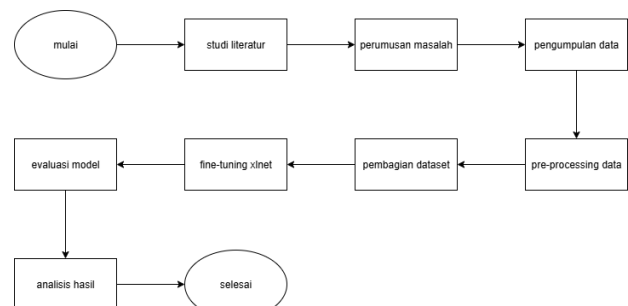
Kemampuan untuk menganalisis dan mendeteksi sarkasme dalam teks media sosial menjadi sangat penting, terutama dalam konteks opini publik. Sarkasme dapat memberikan dampak signifikan terhadap persepsi suatu isu, organisasi, atau individu, sehingga memerlukan pendekatan yang lebih kompleks dibandingkan analisis sentimen konvensional [3], [4]. Hal ini disebabkan oleh kompleksitas linguistik dalam sarkasme, yang sering kali melibatkan penggunaan ironi, metafora, serta elemen budaya yang sulit diinterpretasikan oleh model analisis sentimen berbasis aturan atau metode tradisional [5].

Teknologi pemrosesan bahasa alami (Natural Language Processing/NLP) berbasis deep learning, khususnya model transformer seperti XLNet, memberikan solusi terhadap tantangan ini. XLNet memiliki kemampuan memahami konteks bahasa secara lebih mendalam melalui pendekatan autoregressive dan penggunaan permutasi input yang fleksibel [3]. Dengan mekanisme ini, XLNet mampu menangkap hubungan semantik dan kontekstual yang diperlukan dalam mendeteksi sarkasme, terutama dalam bahasa Indonesia yang memiliki struktur dan karakteristik unik [6].

Penelitian terkait deteksi sarkasme dalam bahasa Indonesia masih terbatas. Sebagian besar penelitian yang ada berfokus pada bahasa Inggris, sementara analisis sarkasme dalam bahasa Indonesia menghadapi tantangan seperti keterbatasan dataset dan kurangnya model yang dioptimalkan untuk memahami karakteristik bahasa tersebut [7], [8]. Untuk mengatasi keterbatasan tersebut, berbagai model transformer yang diadaptasi untuk bahasa Indonesia, seperti IndoBERT, mulai digunakan karena kemampuannya dalam menangkap struktur dan nuansa linguistik lokal secara lebih mendalam. Penelitian sebelumnya menunjukkan bahwa IndoBERT mampu memberikan performa tinggi dalam tugas analisis sentimen berbahasa Indonesia, misalnya pada analisis tweet menjelang Pemilu 2024 [6]. Oleh karena itu, penelitian ini bertujuan mengisi celah tersebut dengan menerapkan XLNet dalam analisis sentimen sarkasme di X berbahasa Indonesia, sehingga diharapkan dapat memberikan kontribusi akademis dalam bidang NLP sekaligus manfaat praktis bagi berbagai pihak, seperti perusahaan, pemerintah, dan organisasi lainnya, dalam memahami opini publik secara lebih akurat dan mendalam [9].

II. METODE PENELITIAN

A. Alur Penelitian



Gbr. 1 Alur Penelitian

Alur penelitian ini secara sistematis, dimulai dengan studi literatur, di mana dilakukan identifikasi topik penelitian, pengumpulan referensi, analisis, dan identifikasi celah penelitian. Tahapan ini membantu menemukan kesenjangan dalam penelitian terdahulu dan menjadi dasar untuk perumusan masalah, yaitu menetapkan tujuan dan fokus utama penelitian berdasarkan celah yang ditemukan. Alur lengkap metode penelitian ditunjukkan pada Gbr. 1.

B. Studi Literatur

Studi literatur dilakukan dengan meninjau penelitian terdahulu terkait deteksi sarkasme dan analisis sentimen pada media sosial. Referensi yang dianalisis mencakup pendekatan machine learning klasik, deep learning, serta model berbasis *Transformer*. Kajian literatur tersebut menunjukkan bahwa sebagian besar penelitian sebelumnya difokuskan pada bahasa Inggris dan masih terbatas pada dataset umum tanpa memperhatikan fenomena linguistik spesifik seperti sarkasme. Selain itu, studi berbasis *Transformer* pada bahasa Indonesia lebih banyak berfokus pada sentimen konvensional, sementara penerapannya pada deteksi sarkasme masih jarang ditemukan. Oleh karena itu, penelitian ini berkontribusi dengan mengevaluasi model *XLNet* untuk mendeteksi sarkasme pada teks media sosial berbahasa Indonesia.

C. Perumusan Masalah

Permasalahan dalam penelitian ini dirumuskan dari temuan studi literatur, yaitu adanya tantangan dalam mendeteksi sarkasme pada analisis sentimen teks media sosial berbahasa Indonesia. Sarkasme bersifat ambigu dan sering bertolak belakang antara makna literal dan maksud penutur sehingga model sentimen konvensional kerap melakukan kesalahan klasifikasi. Tantangan ini menjadi relevan karena analisis sentimen pada media sosial memiliki dampak dalam berbagai bidang seperti pemasaran, layanan pelanggan, dan politik, di mana pemahaman opini publik berpengaruh terhadap pengambilan keputusan. Berdasarkan hal tersebut, penelitian ini memfokuskan permasalahan pada bagaimana memanfaatkan pendekatan berbasis *Transformer*, khususnya *XLNet*, untuk meningkatkan kemampuan sistem dalam memahami konteks linguistik yang kompleks guna mengidentifikasi sarkasme secara lebih akurat.

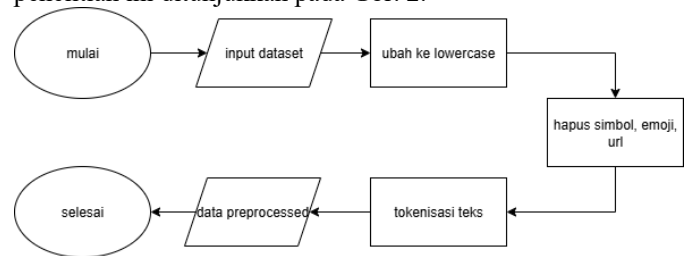
D. Pengumpulan Data

Pada tahap pengumpulan data, penelitian ini menggunakan dataset yang tersedia pada platform *Hugging Face* yang berasal dari platform X (Twitter) dan dikumpulkan melalui API resmi. Dataset tersebut mencakup *tweet* yang dihimpun pada periode Maret 2013 hingga Februari 2020 dengan total lebih dari satu juta *tweet* yang mengandung hashtag relevan seperti #sarcasm, #banjir (“flood”), #bodoh (“stupid”), #anies, dan lainnya. Setelah melalui tahap pembersihan dan penyaringan data berdasarkan relevansi, tersisa 17.838 *tweet* dengan 4.350 *tweet* teridentifikasi sebagai sarkasme (Khotijah et al., 2020). Dataset ini dikenal dalam komunitas penelitian sebagai *Twitter Indonesia Sarcastic* dan telah digunakan secara luas sebagai

sumber data untuk pengembangan dan evaluasi model NLP berbahasa Indonesia.

E. Preprocessing Data

Proses *preprocessing* diperlukan untuk memastikan kualitas data dan menyiapkan dataset agar dapat digunakan pada tahap pemodelan. Tahapan ini penting karena teks pada media sosial umumnya mengandung banyak elemen yang tidak relevan seperti emoji, simbol, tautan, serta variasi penulisan informal. Pada penelitian ini *preprocessing* dilakukan untuk menormalkan struktur teks, meningkatkan konsistensi linguistik, serta mempermudah proses tokenisasi oleh model *Transformer*. Tahapan *preprocessing* yang digunakan dalam penelitian ini ditunjukkan pada Gbr. 2.



Gbr. 2 Alur Preprocessing Data

1. Input Dataset

Dataset mentah hasil *crawling* Twitter dimasukkan sebagai data awal untuk diproses.

2. Lowercasing

Seluruh teks dikonversi ke huruf kecil untuk menyamakan representasi kata.

3. Noise Removal

Elemen non-linguistik seperti URL, simbol, emoji, dan karakter tidak relevan dihapus dari teks.

4. Tokenisasi

Teks dibagi menjadi token agar dapat diolah oleh sistem dan model.

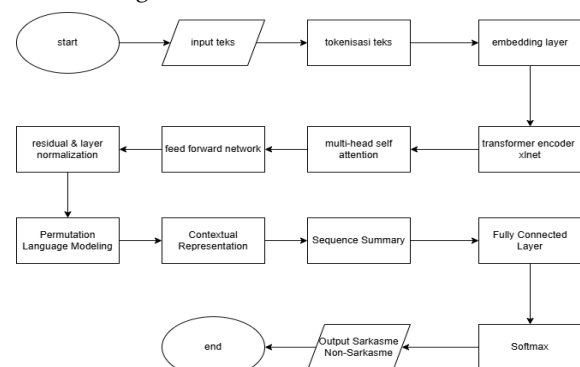
5. Output Data Preprocessed

Hasil akhir berupa teks yang telah dibersihkan dan siap digunakan pada tahap pemodelan.

F. Pembagian Data

Pembagian data merupakan langkah penting setelah proses *preprocessing*, di mana data yang telah dipersiapkan dibagi menjadi 3 bagian utama dengan beberapa skenario untuk mengevaluasi performa model secara lebih menyeluruh. Training set digunakan untuk melatih model, validation set untuk memantau performa selama pelatihan dan mencegah *overfitting*, sedangkan test set digunakan untuk mengevaluasi performa akhir model pada data yang belum pernah dilihat.

G. Fine-Tuning XLNet



Gbr. 3 Fine-tuning XLNet

Alur lengkap *fine-tuning XLNet* ditunjukkan pada Gbr. 3. Arsitektur *XLNet* dibangun di atas konsep *Transformer*, sehingga mekanisme utama yang digunakan adalah *Self-attention* dan *Feed Forward Network* sebagaimana pada *Transformer Encoder*. Proses diawali dengan tokenisasi teks menggunakan *XLNet Tokenizer*, kemudian token direpresentasikan dalam bentuk *Embedding*. *Embedding* tersebut diproses oleh lapisan *Transformer Encoder* yang terdiri dari *Multi-head Self-attention*, *Feed Forward Network*, serta *Residual Connection* dan *Layer Normalization*. Perbedaan utama *XLNet* dengan *Transformer* konvensional terletak pada penerapan *Permutation Language Modeling*, yang memungkinkan model mempelajari ketergantungan kontekstual secara dua arah tanpa *Masking*. Representasi kontekstual yang dihasilkan kemudian dirangkum melalui *Sequence Summary Layer* dan diteruskan ke lapisan klasifikasi untuk menghasilkan prediksi kelas sarkasme dan non-sarkasme.

H. Skenario Uji Coba

Pada scenario uji coba, dataset dibagi menjadi data training, validasi, dan testing dengan tiga skenario proporsi, yaitu 70:15:15, 60:20:20, dan 80:10:10 untuk mengevaluasi pengaruh ukuran data pelatihan terhadap performa model *XLNet* dalam mendeteksi sarkasme. Model yang digunakan merupakan *pre-trained XLNet* (indonesian-pretrained) dengan lapisan keluaran berupa *softmax* untuk klasifikasi multi-kelas untuk klasifikasi biner. Beberapa parameter pelatihan dikonfigurasi untuk eksplorasi performa, antara lain variasi *dropout* (0.1 dan 0.3), *learning rate* ($3e-5$ dan $5e-5$), *batch size* (16 dan 32), jumlah *epoch* (30 dan 50), *sequence length* (256 dan 512), *weight decay* (0.01), serta penggunaan *optimizer* AdamW dan fungsi *loss cross-entropy*. Skenario ini disusun untuk menemukan konfigurasi yang menghasilkan performa terbaik, yang ditunjukkan pada Tabel 1.

TABEL I
HYPERPARAMETER TUNING

Tahap	Parameter	Nilai
Pembagian Data	Data training, data validasi, data testing	70, 15, dan 15 60, 20, dan 20 80, 10, dan 10
Desain Model	<i>XLNet Pre-trained</i>	<i>xlnet-base-cased</i>
Training Model	<i>Learning Rate</i>	$3e-5$, $5e-5$
	<i>Batch Size</i>	16, 32
	<i>Dropout</i>	0.1, 0.3
	<i>Optimizer</i>	AdamW

	<i>Loss Function</i>	<i>Cross-Entropy Loss</i>
	Epoch	30 sd 50 (<i>early stopping</i>)

I. Evaluasi Model

Setelah proses pelatihan selesai, performa model dievaluasi menggunakan data testing yang diperoleh melalui pembagian data secara *stratified* sesuai proporsi masing-masing skenario. Evaluasi dilakukan menggunakan *confusion matrix* serta *classification report* untuk mengamati jenis kesalahan prediksi pada kelas sarkasme dan non-sarkasme. *Confusion matrix* memberikan informasi terkait distribusi prediksi benar dan salah pada masing-masing kelas, sedangkan *classification report* memberikan nilai akurasi dan *macro-F1* sebagai metrik utama untuk menilai kualitas model pada data yang seimbang maupun tidak seimbang. Setelah seluruh metrik diperoleh, dilakukan perbandingan performa pada tiga skenario pembagian data (60:20:20, 70:15:15, dan 80:10:10) untuk menentukan konfigurasi pembagian data yang paling optimal dalam mendeteksi sarkasme pada teks X. Tahap evaluasi ini menjadi dasar dalam penarikan kesimpulan mengenai efektivitas model *XLNet* dalam mengidentifikasi sarkasme pada media sosial.

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Penelitian ini memakai dataset sarkasme berbahasa Indonesia yang asalnya dari Twitter diambil lewat *API* resmi pada Maret 2013 sampai Februari 2020. Dataset tersebut berjumlah 17.838 tweet yang kemudian dibersihkan sehingga tersisa 2.684 tweet yang relevan dan seimbang, dengan 671 tweet di antaranya berlabel sarkasme.

Untuk kemudahan akses, pada penelitian ini penulis menggunakan versi yang dipublikasikan di *Hugging Face* (*wl1wo/twitter_indonesia_sarcastic*).

B. Under Sampling

Setelah proses pengumpulan data, pada sub bab ini disajikan hasil mengenai penerapan metode under sampling karena jumlah antar label belum seimbang.

C. Preprocessing Data

Setelah data dikumpulkan dan disimpan dalam format CSV, dilakukan *preprocessing* ringan untuk memastikan tidak ada yang terlewat sebelum diproses model. Karena dataset yang diambil dari *Hugging Face* sudah bersih, tahapan ini diterapkan sebagai cek dan standarisasi saja yaitu *lowercase*, hapus simbol, emoji, url, dan tokenisasi teks.

D. Pembagian Data

Pada dataset asli dari *Hugging Face*, data sudah terbagi menjadi *train*, *validation*, dan *test*. Tetapi komposisinya tidak seimbang dimana label non-sarkasme jauh lebih banyak

daripada sarkasme pada tiap split. Agar evaluasi lebih adil dan model tidak berat sebelah, penulis melakukan pembagian ulang menggunakan dataset yang telah dilakukan *under sampling* menjadi 3 skenario yaitu 70:15:15, 60:20:20, dan 80:10:10.

E. Fine Tuning XLNet dan Validasi

Setelah data ditokenisasi, model XLNet dibangun menggunakan XLNetConfig dengan dua label untuk klasifikasi biner dan `summary_dropout_prob=0.1`, lalu dimuat melalui `XLNetForSequenceClassification.from_pretrained(...)`. Fungsi evaluasi `compute_metrics` disiapkan untuk menghitung akurasi, *precision*, *recall*, dan *F1-score* sebagai metrik utama selama pelatihan. Tahap berikutnya mendefinisikan parameter pelatihan melalui `TrainingArguments`, di mana evaluasi dan penyimpanan dilakukan setiap *epoch*, serta `load_best_model_at_end=True` dengan acuan metrik *F1*. *Hyperparameter* inti yang digunakan meliputi *learning rate*, *batch size*, *epoch*, dan *weight decay*. Proses pelatihan dilakukan dengan `Trainer` yang menggabungkan model, argumen pelatihan, dataset *tokenized*, *tokenizer*, dan data *collator*, serta dilengkapi `EarlyStoppingCallback` agar pelatihan dihentikan bila performa tidak membaik dalam beberapa *epoch*. Evaluasi dilakukan menggunakan `trainer.train()` dan `trainer.predict(...)`, menghasilkan *classification report* dan *confusion matrix* untuk melihat pola kesalahan model. Model dan *tokenizer* terbaik kemudian disimpan menggunakan `save_pretrained` sehingga dapat digunakan kembali pada tahap pengujian tanpa perlu pelatihan ulang. Hasil training dan validasi adalah sebagai berikut.

1. Skenario 1 (70:15:15)

Pada tahap ini dilakukan pelatihan dan validasi model XLNet menggunakan konfigurasi *hyperparameter* yang telah diuji. Parameter terbaik untuk skenario 1 disajikan pada Tabel II.

TABEL II
PARAMETER KONFIGURASI TERBAIK SKENARIO 1

ID	Epoch	Learning rate	Batch size	Sequence length	Dropout
Exp-17	50	3e-5	16	256	0.1
Exp-18	50	3e-5	16	256	0.3
Exp-19	50	3e-5	16	512	0.1
Exp-20	50	3e-5	16	512	0.3

Secara keseluruhan, keempat eksperimen menghasilkan performa yang relatif serupa dengan akurasi dan F1-score mencapai 0.84.

2. Skenario 2 (60:20:20)

Pada tahap ini dilakukan pelatihan dan validasi model XLNet menggunakan konfigurasi *hyperparameter* yang telah diuji. Parameter terbaik untuk skenario 2 disajikan pada Tabel III.

TABEL III

PARAMETER KONFIGURASI TERBAIK SKENARIO 2

ID	Epoch	Learning rate	Batch size	Sequence length	Dropout
Exp-19	50	3e-5	16	512	0.1
Exp-20	50	3e-5	16	512	0.3

Secara keseluruhan, kedua eksperimen menghasilkan performa yang relatif serupa dengan akurasi dan F1-score mencapai 0.81.

3. Skenario 3 (80:10:10)

Pada tahap ini dilakukan pelatihan dan validasi model XLNet menggunakan konfigurasi *hyperparameter* yang telah diuji. Parameter terbaik untuk skenario 3 disajikan pada Tabel IV.

TABEL IV

PARAMETER KONFIGURASI TERBAIK SKENARIO 1

ID	Epoch	Learning rate	Batch size	Sequence length	Dropout
Exp-1	30	3e-5	16	256	0.1
Exp-2	30	3e-5	16	256	0.3
Exp-3	30	3e-5	16	512	0.1
Exp-4	30	3e-5	16	512	0.3

Secara keseluruhan, keempat eksperimen menghasilkan performa yang relatif serupa dengan akurasi dan F1-score mencapai 0.81.

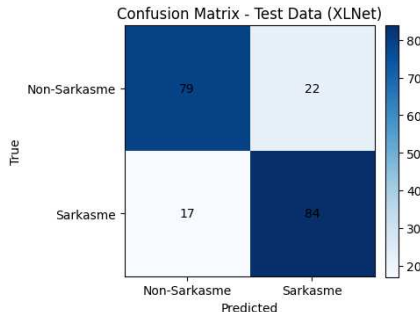
F. Evaluasi Model

Berdasarkan hasil analisis seluruh percobaan, model XLNet pada skenario 70:15:15 eksperimen ke-17 dipilih sebagai model terbaik karena menghasilkan skor validasi tertinggi atau mendekati tertinggi secara konsisten, menggunakan konfigurasi *hyperparameter* yang lebih ringan dibanding skenario lain, serta mencapai *macro-F1* yang baik dan seimbang pada kedua label. Model juga menunjukkan kemampuan menangkap sarkasme dengan *recall* yang tinggi. Model terpilih kemudian dievaluasi pada data testing yang terpisah tanpa penalaan ulang (*no re-tuning*), dan hasilnya disajikan melalui *classification report* dan *confusion matrix* untuk memperlihatkan metrik per label (*precision*, *recall*, dan *F1-score*) serta distribusi prediksi (TP, FP, FN, TN). Evaluasi ini menjadi dasar penyusunan pembahasan dan penarikan kesimpulan akhir penelitian. Hasil dari evaluasi model disajikan pada Tabel V.

TABEL V
HASIL EVALUASI

		Performa	
		F1-Score	Akurasi
XLNet	Skenario 1 (Exp-17)	0.80	0.81

Hasil evaluasi model terbaik pada skenario 1 dirangkum pada Tabel V yang menunjukkan akurasi sebesar 0.81.



Gbr. 4 Confusion Matrix

Confusion Matrix XLNet pada Gbr. 4, dimana dari 202 tweet diperoleh $TN = 79$ (non-sarkasme terdeteksi benar), $TP = 84$ (sarkasme terdeteksi benar), $FP = 22$ (non-sarkasme salah terklasifikasi sebagai sarkasme), dan $FN = 17$ (sarkasme terlewat). Pola ini menunjukkan bahwa *XLNet* memiliki kemampuan deteksi yang baik terhadap sarkasme, meskipun masih terdapat sebagian kasus non-sarkasme yang diklasifikasikan sebagai sarkasme.

	precision	recall	f1-score	support
Non-Sarkasme	0.82	0.78	0.80	101
Sarkasme	0.79	0.83	0.81	101
accuracy			0.81	202
macro avg	0.81	0.81	0.81	202
weighted avg	0.81	0.81	0.81	202

Gbr. 5 Classification Report

Classification Report XLNet pada Gbr. 5 menunjukkan akurasi sebesar 0.81 dan macro-F1 sebesar 0.81. Secara per label, kelas non-sarkasme memiliki *precision* 0.82, *recall* 0.78, dan *F1-score* 0.80, sedangkan kelas sarkasme memiliki *precision* 0.79, *recall* 0.83, dan *F1-score* 0.81. Hasil tersebut menunjukkan bahwa model memiliki performa yang seimbang pada kedua kelas dan cukup sensitif dalam mendeteksi sarkasme.

IV. KESIMPULAN

Berdasarkan hasil penelitian, model *XLNet* berhasil membentuk pipeline deteksi sarkasme pada teks berbahasa Indonesia melalui tahapan *preprocessing* dan tokenisasi yang memungkinkan model menangkap konteks secara lebih efektif. Model menunjukkan kemampuan klasifikasi yang baik dalam membedakan sarkasme dan non-sarkasme, dengan performa yang lebih stabil pada label sarkasme yang umumnya sulit dikenali dalam analisis sentimen konvensional. Selain itu, eksperimen *fine-tuning* menunjukkan bahwa pemilihan

hyperparameter berpengaruh terhadap performa model, di mana *learning rate* yang lebih kecil memberikan pelatihan yang stabil, *batch size* yang moderat membantu mencegah *overfitting*, dan *sequence length* antara 256–512 memberikan konteks yang cukup untuk mendukung peningkatan nilai F1.

V. SARAN

Berdasarkan kesimpulan dan keterbatasan penelitian, saran yang dapat diberikan adalah sebagai berikut:

1. Penelitian selanjutnya dapat mencoba untuk memperluas data, jangan hanya mengambil dari Twitter atau X. Tambahkan dari Instagram, TikTok atau sosial media yang lain dan menentukan periode waktu yang lebih beragam agar model lebih kuat di berbagai bentuk gaya bahasa.
2. Penelitian selanjutnya dapat mencoba model transformer lain seperti *IndoRoBERTa*, *RoBERTa*, *DistilBERT* atau *hybrid head* seperti *BiGRU* atau *BiLSTM* di atas *embedding XLNet* untuk membandingkan akurasi dan efisiensi, terutama pada label sarkasme yang sulit dipahami.

UCAPAN TERIMA KASIH

Peneliti menyampaikan puji syukur kepada Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga penelitian dan penulisan artikel ilmiah ini dapat diselesaikan dengan baik. Peneliti juga mengucapkan terima kasih kepada Dosen Pembimbing dan Koordinator Program Studi atas bimbingan, arahan, dan masukan yang diberikan selama proses penelitian berlangsung. Ucapan terima kasih turut disampaikan kepada orang tua, keluarga, serta rekan-rekan di Jurusan Teknik Informatika angkatan 2021 yang senantiasa memberikan dukungan, bantuan, dan motivasi hingga penelitian ini dapat diselesaikan dengan lancar.

REFERENSI

- [1] S. Farabi, T. Ranasinghe, D. Kanojia, Y. Kong, and M. Zampieri, "A Survey of Multimodal Sarcasm Detection," pp. 8020–8028, 2024, doi: 10.24963/ijcai.2024/887.
- [2] D. Suhartono, W. Wongso, and A. Tri Handoyo, "IdSarcasm: Benchmarking and Evaluating Language Models for Indonesian Sarcasm Detection," *IEEE Access*, vol. 12, pp. 87323–87332, 2024, doi: 10.1109/ACCESS.2024.3416955.
- [3] K. T. Habib. A, Afridi. Y. S, Sher. M, "Relevance Classification of Flood-Related Tweets Using XLNET Deep Learning Model," no. May, pp. 158–164, 2024.
- [4] D. Olaniyan, R. O. Ogundokun, O. P. Bernard, J. Olaniyan, R. Maskeliūnas, and H. B. Akande, "Utilizing an Attention-Based LSTM Model for Detecting Sarcasm and Irony in Social Media," *Computers*, vol. 12, no. 11, pp. 1–15, 2023, doi: 10.3390/computers12110231.
- [5] A. Kumar J, T. Esther Trueman, and E. Cambria, "Fake News Detection Using XLNet Fine-Tuning Model," *2021 Int. Conf. Comput. Intell. Comput. Appl. ICCICA 2021*, no. July 2022, 2021, doi: 10.1109/ICCICA52458.2021.9697269.
- [6] L. Geni, E. Yulianti, and D. I. Sensus, "Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using IndoBERT Language Models," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 9, no. 3, pp. 746–757, 2023, doi: 10.26555/jiteki.v9i3.26490.
- [7] Y. Nurdiansyah, F. Rahman, and P. Pandunata, "Analisis Sentimen Opini Publik Terhadap Undang-Undang Cipta Kerja pada Twitter Menggunakan Metode Naive Bayes Classifier," *Pros. Semin. Nas. Sains Teknol. dan Inov. Indones.*, vol. 3, no. November, pp. 201–212,

- [8] 2021, doi: 10.54706/senastindo.v3.2021.158.
B. S. Wiguna, C. V. Hudyanti, A. Alqis Rausanfita, and A. Z. Arifin, "Sarcasm Detection Engine for Twitter Sentiment Analysis using Textual and Emoji Feature," *J. Ilmu Komput. dan Inf.*, vol. 14, no. 1, pp. 1–8, 2021, doi: 10.21609/jiki.v14i1.812.
- [9] A. Palaniammal and P. Anandababu, "Sarcasm detection on social data: heuristic search and deep learning," *LAES Int. J. Artif. Intell.*, vol. 13, no. 4, pp. 4695–4702, 2024, doi: 10.11591/ijai.v13.i4.pp4695-4702.