ISSN: 3110-0864 (Media Online)
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | pp: 232-240
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.146

Deteksi Sarkasme Implisit pada Teks Berbahasa Indonesia Menggunakan Hybrid IndoBERT dan Knowledge Graph

Novia Ardani[1], Yakiya Ratna Sari[2*], Riska Dwi Handayani[3]

[1,2,3]Teknik Informatika, STMIK Bina Patria, Kota Magelang, Indonesia
Email: [1]viaani863@gmail.com, [2*]ratnasariyakia@email.com, [3]riska@stmikbinapatria.ac.id
(* Email Corresponding Author: ratnasariyakia@email.com)
Dikirim: 25 April 2026 | Diterima: 30 Juni 2026 | Dipublish: 4 Juli 2026

Abstrak

Sarkasme implisit merupakan bentuk ungkapan sindiran yang tidak ditandai secara langsung oleh kata atau frasa sarkastik, sehingga menjadi tantangan dalam proses deteksi otomatis menggunakan teknik pemrosesan bahasa alami. Penelitian ini mengembangkan pendekatan hibrida yang mengombinasikan model bahasa *IndoBERT* dengan *Knowledge Graph* yang dibangun berdasarkan leksikon bahasa Indonesia. *Knowledge Graph* tersebut terdiri atas empat jenis pengetahuan, yaitu leksikon sentimen, indikator sarkasme, pasangan kontradiksi semantik, serta kamus normalisasi bahasa slang. Representasi kontekstual yang dihasilkan oleh *IndoBERT* kemudian diperkaya melalui mekanisme *multi-head attention* yang memanfaatkan fitur dari *Knowledge Graph*, selanjutnya digabungkan menggunakan lapisan fusion sebelum dilakukan proses klasifikasi. Dataset yang digunakan adalah Reddit Indonesia Sarcastic yang tersedia secara publik melalui *HuggingFace Datasets*. Data dibagi menjadi 70% untuk pelatihan, 15% untuk validasi, dan 15% untuk pengujian. Pelatihan model dilakukan menggunakan optimizer AdamW dengan *learning rate* sebesar 2×10^{-5} selama lima epoch. Hasil pengujian menunjukkan bahwa model yang diusulkan memperoleh akurasi sebesar 79,65%, F1-score sebesar 80,31%, dan ROC-AUC sebesar 85,46%. Temuan ini menunjukkan bahwa pendekatan hibrida mampu meningkatkan kemampuan model dalam mengenali kontradiksi semantik yang menjadi karakteristik utama sarkasme implisit dalam bahasa Indonesia. Penelitian ini memberikan kontribusi terhadap pengembangan sistem deteksi sarkasme berbahasa Indonesia yang lebih efektif melalui integrasi pengetahuan leksikal berbasis graf.

Kata Kunci: Sarkasme Implisit, IndoBERT, Knowledge Graph, Klasifikasi Teks, NLP Bahasa Indonesia

Abstract

Implicit sarcasm is a form of figurative expression in which sarcastic intent is conveyed without explicit sarcasm markers, making it challenging for natural language processing systems to detect. This study proposes a hybrid approach that integrates the IndoBERT language model with a lexicon-based Knowledge Graph designed for the Indonesian language. The constructed Knowledge Graph incorporates four primary knowledge sources, namely sentiment lexicons, sarcasm indicators, semantic contradiction pairs, and a slang normalization dictionary. Contextual semantic representations generated by IndoBERT are enriched through a multi-head attention mechanism that leverages Knowledge Graph features, followed by a fusion layer before the final classification stage. The experiments were conducted using the publicly available Reddit Indonesia Sarcastic dataset from HuggingFace Datasets. The dataset was partitioned into 70% training data, 15% validation data, and 15% testing data. Model training employed the AdamW optimizer with a learning rate of 2×10^{-5} for five epochs. Experimental results demonstrate that the proposed model achieved an accuracy of 79.65%, an F1-score of 80.31%, and a ROC-AUC of 85.46% on the test set. These findings indicate that the hybrid approach effectively captures semantic contradiction patterns, which are a defining characteristic of implicit sarcasm in Indonesian text. This research contributes to the development of more accurate Indonesian sarcasm detection systems by incorporating domain-specific lexical knowledge through graph-based representations.

***Keywords**: Implicit Sarcasm, IndoBERT, Knowledge Graph, Text Classification, Indonesian NLP*

1. PENDAHULUAN

Platform media sosial seperti Reddit telah menjadi wadah ekspresi yang luas bagi masyarakat Indonesia. Beragam jenis pendapat, sindiran, dan emosi diungkapkan melalui tulisan singkat yang penuh makna tersirat. Salah satu bentuk ungkapan yang sering muncul tetapi sulit dimengerti oleh sistem kecerdasan buatan adalah sarkasme. Sarkasme adalah jenis ironi verbal yang di mana arti yang sebenarnya bertentangan dengan arti harfiah dari kata-kata yang dipakai. Tantangan ini menjadi semakin rumit saat sarkasme bersifat implisit, yaitu tidak memanfaatkan penanda sarkasme yang jelas seperti kata "seharusnya" atau tanda seru berlebihan, melainkan bergantung pada konteks, pengetahuan budaya, dan pola kontradiksi semantik yang halus [1].

Mendeteksi sarkasme dalam teks berbahasa Indonesia lebih sulit daripada dalam bahasa Inggris. Terdapat beberapa faktor yang menyebabkan hal ini, seperti variabilitas penggunaan bahasa informal dan slang, minimnya sumber daya NLP *(Natural Language Processing)* untuk bahasa Indonesia, serta pengaruh budaya

This Journal is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

ISSN: 3110-0864 (Media Online)
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | pp: 232-240
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.146

lokal yang kuat dalam membentuk arti sarkastik [2]. Sistem deteksi sentimen tradisional yang hanya bergantung pada polaritas kata biasanya tidak berhasil dalam mengidentifikasi sarkasme tersirat karena sarkasme sering kali menggunakan kata-kata positif dalam konteks yang memiliki makna negatif. Berbagai penelitian telah berupaya menyelesaikan permasalahan deteksi sarkasme. Zhang dkk. [3] mengusulkan pendekatan berbasis *transformer* dengan mekanisme *multi-head attention* untuk deteksi sarkasme pada dataset *Self-Annotated Reddit Corpus* (SARC) yang bersumber dari platform Reddit dan memperoleh *precision, recall,* serta *F1-score* masing-masing sebesar 0,917.

Namun penelitian ini berfokus pada teks berbahasa Inggris dan belum mempertimbangkan karakteristik linguistik bahasa Indonesia seperti penggunaan *slang* dan konteks budaya lokal. Selanjutnya, Potamias dkk. [4] memanfaatkan model berbasis *transformer* dikombinasikan dengan *recurrent convolutional neural network* untuk deteksi sarkasme dan menunjukkan bahwa representasi kontekstual berbasis *pre-trained model* mampu meningkatkan performa secara signifikan dibandingkan model sekuensial tradisional. Di sisi lain, Wallace dkk. [5] mengusulkan model berbasis ELMo dengan representasi kata tingkat karakter untuk mendeteksi sarkasme pada berbagai dataset termasuk Reddit dan Twitter, dan mencapai performa *state-of-the-art* pada enam dari tujuh dataset yang diuji, namun pendekatan ini belum mengintegrasikan pengetahuan leksikal eksternal yang spesifik terhadap suatu bahasa.

Untuk bahasa Indonesia secara khusus, Saputra dkk. [6]mengembangkan *IndoBERT-Sentiment,* sebuah pengklasifikasi sentimen berbasis IndoBERT yang mempertimbangkan konteks topikal teks berbahasa Indonesia dan memperoleh *F1-score* sebesar 0,856. Namun penelitian ini tidak secara khusus menangani sarkasme implisit. Penelitian Willie dkk. [7] memperkenalkan *IndoNLU Benchmark* yang mencakup berbagai tugas NLP bahasa Indonesia termasuk analisis sentimen, tetapi deteksi sarkasme implisit belum termasuk dalam cakupan evaluasinya sehingga masih terdapat *gap* yang perlu diteliti lebih lanjut.

Berdasarkan penelitian sebelumnya, terdapat kelemahan yang signifikan dalam deteksi sarkasme implisit berbahasa Indonesia, terutama terkait dengan minimnya integrasi pengetahuan leksikal domain-spesifik bahasa Indonesia dalam model *deep learning*, kurangnya fokus pada pola kontradiksi semantik sebagai indikasi sarkasme implisit, dan belum digunakannya *Knowledge Graph* untuk memperkaya konteks dalam deteksi sarkasme berbahasa Indonesia.

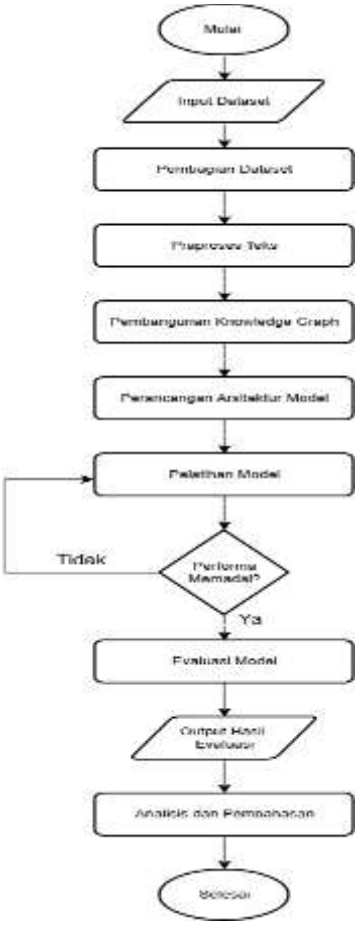
Studi ini bertujuan untuk menciptakan sistem deteksi sarkasme tersembunyi dalam teks berbahasa Indonesia dengan mengajukan arsitektur hibrida yang mengintegrasikan *IndoBERT* sebagai model bahasa kontekstual utama dengan *Knowledge Graph* yang didasarkan pada leksikon bahasa Indonesia. Kontribusi utama dari penelitian ini mencakup desain arsitektur *Hybrid IndoBERT-KG* dengan mekanisme *KG Attention* untuk memperkaya representasi semantik, pengembangan *Knowledge Graph* berbasis leksikon bahasa Indonesia yang meliputi leksikon sentimen, penanda sarkasme, pasangan kontradiksi, dan kamus slang, serta evaluasi menyeluruh pada dataset *Reddit Indonesia Sarcastic* dengan menggunakan metrik akurasi, F1-Score, dan ROC-AUC.

2. METODOLOGI PENELITIAN

Studi ini terdiri atas empat langkah utama, yaitu pengumpulan dan praproses data, pengembangan Knowledge Graph, perancangan arsitektur model hibrida, serta pelaksanaan pelatihan dan evaluasi. Ilustrasi keseluruhan alur penelitian ditampilkan dalam Gambar 1.

This Journal is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

ISSN: 3110-0864 (Media Online)
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | pp: 232-240
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.146



Gambar 1. Flowchart Penelitian

2.1 Pengumpulan Data

Studi ini memanfaatkan dataset Reddit Indonesia Sarcastic yang dapat diakses secara publik melalui platform *HuggingFace Datasets* dengan nama *w11wo/reddit_indonesia_sarcastic* [8]. Dataset ini merupakan kumpulan teks dalam bahasa Indonesia yang diambil dari platform Reddit dan telah diberi anotasi secara manual dengan label biner, yaitu sarkasme (1) dan non-sarkasme (0).

Dataset dibagi menggunakan strategi *stratified split* untuk mempertahankan proporsi kelas pada setiap subset data, dengan rasio 70% data latih, 15% data validasi, dan 15% data uji sebagaimana ditunjukkan pada Tabel 1.

Tabel 1. Distribusi Dataset

Label	Train	Validasi	Test	Total
Sarcasm	470	100	101	671
Non-Sarcasm	1409	302	302	2013
Total	1879	402	403	2684

Tahap praproses meliputi dua proses utama. Pertama, normalisasi slang dilakukan menggunakan kamus yang dibangun secara manual dan berisi lebih dari 60 pasangan kata informal ke bentuk formal dalam bahasa Indonesia, seperti ditunjukkan pada Tabel 2. Kedua, tokenisasi dilakukan menggunakan tokenizer IndoBERT dengan panjang maksimum 128 token, serta menerapkan *padding* dan *truncation* untuk menyeragamkan dimensi input.

Tabel 2. Contoh Normalisasi Slang

Slang	Formal
gak/ga/nggak	tidak
bgt / banget	sangat
wkwk / haha	[tawa]
emang / emng	memang
udah / udh	sudah
gimana / gmn	bagaimana

Copyright © 2026 **Author**, Page 234
This Journal is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

ISSN: 3110-0864 (Media Online)
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | pp: 232-240
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.146

2.2 Pembangunan Knowledge Graph

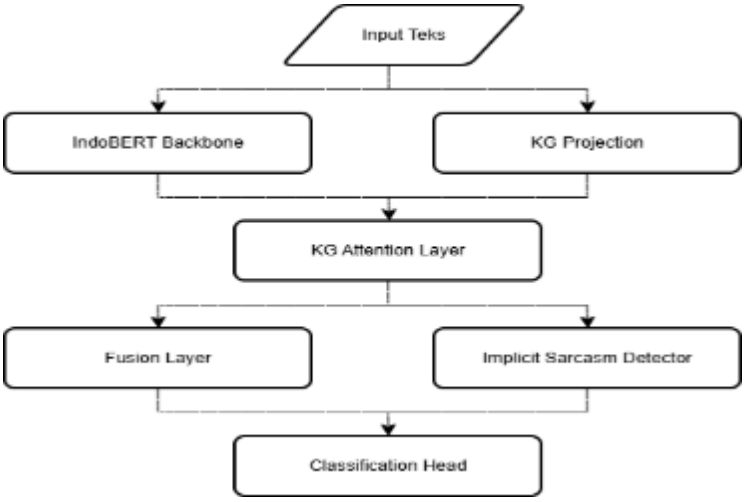
Knowledge Graph (KG) dibangun secara manual berdasarkan leksikon dan pola linguistik bahasa Indonesia [9]. KG berperan sebagai sumber pengetahuan eksternal yang menyediakan informasi semantik tambahan yang tidak selalu dapat dipahami secara optimal oleh model bahasa seperti IndoBERT, khususnya pada kasus sarkasme implisit yang bergantung pada konteks budaya dan pola kontradiksi dalam bahasa indonesia [10]. KG yang dikembangkan terdiri dari 4 komponen utama:

a. Leksikon Sentimen, yang memuat 45 kata bersentimen positif (misalnya *bagus, hebat, sukses*) dan 49 katabersentimen negatif (misalnya *buruk, gagal, kacau*) dalam bahasa Indonesia.
b. Penanda Sarkasme, yang berisi 27 frasa yang umum digunakan dalam ungkapan sarkastik bahasa Indonesia, seperti "oh tentu", "pasti dong", "sungguh luar biasa", dan "betapa beruntungnya".
c. Pasangan Kontraksi Semantik, yang terdiri dari 19 pasangan kata dengan makna yang saling bertentangan, seperti *(bagus, parah), (berhasil, gagal),* dan *(pintar, bodoh).* Kemunculan pasangan kata tersebut dalam satu kalimat dapat menjadi indikator sarkasme implisit.
d. Kamus Slang, yang memuat lebih dari 60 entri bahasa gaul Indonesia beserta bentuk formalnya dan digunakan pada tahap praproses data.

Berdasarkan komponen tersebut, fungsi *extract_kg_features()* menghasilkan vektor fitur berdimensi 128 untuk setiap teks. Fitur yang dihasilkan meliputi skor sentimen, kepadatan penanda sarkasme, skor kontradiksi semantik, rasio penggunaan slang, serta kepadatan tanda baca ekspresif seperti tanda seru, tanda tanya, dan elipsis.

2.3 Arsitektur Model Hybrid IndoBERT-KG

Model yang diusulkan adalah *hybrid* yang menggabungkan *IndoBERT* dengan KG melalui mekanisme *attention.* Arsitektur lengkapnya ditunjukkan pada Gambar 2.



Gambar 2. Arsitektur Model Hybrid IndoBERT-KG

Arsitektur terdiri dari enam komponen utama yaitu :

a. IndoBERT Backbone
Model *IndoBERT-base-p1* [2]digunakan sebagai *encoder* teks. Token (CLS) pada lapisan terakhir sebagai representasi semantik kontekstual berdimensi 768.
b. KG Projection
Vektor fitur KG berdimensi 128 diproyeksikan ke dimensi 768 melalui dua lapisan *fully connected* dengan aktivasi ReLU dan *dropout* 0,3 agar kompatibel dengan dimensi representasi BERT.
c. KG Attention Layer
Mekanisme *multi-head attention* dengan 4 head digunakan untuk memperkaya representasi IndoBERT menggunakan pengetahuan dari KG bertindak sebagai *key* dan *value.* Output dinormalisasi menggunakan *LayerNorm* dengan *residual connection* [11]. Persamaan *attention* yang digunakan adalah :

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}}) \cdot V \quad (1)$$

d. Fusion Layer

Copyright © 2026 **Author**, Page 235
This Journal is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

ISSN: 3110-0864 (Media Online)
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | pp: 232-240
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.146

Representasi yang telah diperkaya KG digabungkan dengan representasi KG melalui konkatenasi, kemudian dikompresi menggunakan lapisan *fully connected*+*LayerNorm*+aktivasi GELU+*dropout* [12].

e. Implicit Sarcasm Detector
Sub-jaringan khusus berdimensi 768→256→64 dengan aktivasi ReLU yang dirancang untuk menangkap pola sarkasme implisit dari representasi yang telah diperkaya KG.

f. Classification Head
Lapisan *fully connected* terakhir (832→128→2) menghasilkan probabilitas kelas sarkasme dan non-sarkasme.

2.4 Prosedur Pelatihan

Proses Pelatihan model dilakukan menggunakan optimizer AdamW [13]dengan konfigurasi sebagaimana ditunjukkan pada Tabel 3. *Weight decay* diterapkan pada seluruh parameter kecuali bias dan *LayerNorm* sebagai bentuk regulasi.Selain itu, digunakan *linearwarmup scheduler* dengan rasio *warmup* sebesar 0,1 dari total langkah pelatihan. Fungsi *loss* yang digunakan adalah *CrossEntropyLoss* dengan *class weighting* untuk mengatasi ketidakseimbangan kelas pada dataset.

Tabel 3. Konfigurasi Eksperimen

Parameter	Nilai
Model Dasar	indobenchmark/indobert-base-p1
Panjang Maksimum Token	128
Batch Size	16
Epoch	5
Learning Rate	2×10^{-5}
Warmup Ratio	0,1
Weight Decay	0,01
Optimizer	AdamW
KG Embedding Dimension	128
KG Attention Heads	4
Dropout	0,3
Aktivitasi Fusion	GELU

3. HASIL DAN PEMBAHASAN

3.1 Hasil Evaluasi Model

Evaluasi model dilakukan pada data uji yang belum pernah dilihat selama proses pelatihan. Tabel 4 menunjukkan hasil evaluasi model *Hybrid IndoBERT-KG* yang diusulkan berdasarkan metrik *precision, recall, F1-score, accuracy,* dan ROC-AUC [14].

Tabel 4. Hasil Evaluasi pada Data Uji

Kelas	Precision	Recall	F1-Score	Support
Non-Sarcasm	89,57%	82,45%	85,86%	302
Sarcasm	57,60%	71,29%	63,72%	101
Accuracy			79,65%	403
Macro Avg	73,58%	76,87%	74,79%	403
Weighted Avg	81,56%	79,65%	80,31%	403
ROC-AUC			85,46%	

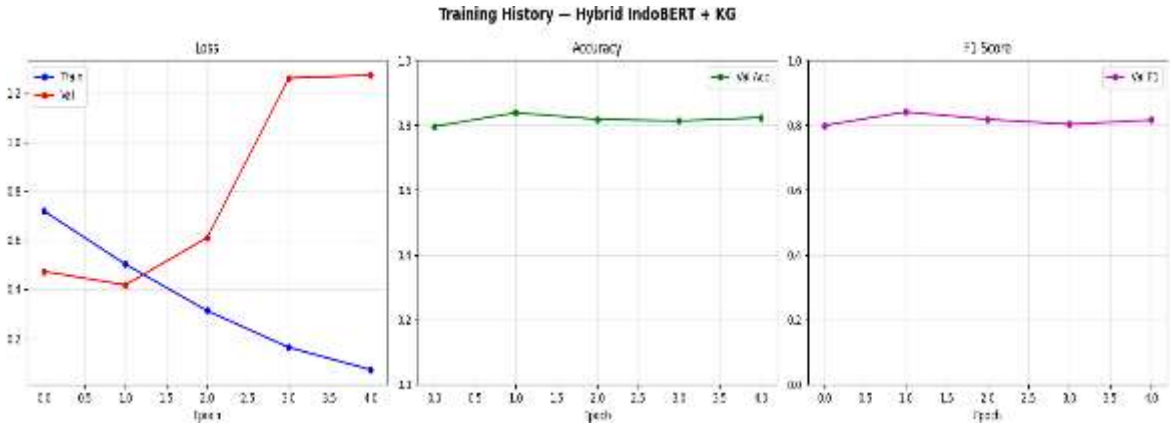
Model yang diusulkan mencapai akurasi sebesar 79,65% dan *F1-score weighted* sebesar 80,31% pada data uji. Nilai ROC-AUC sebesar 85,46% menunjukkan bahwa model memiliki kemampuan diskriminasi nilai yang baik antara kelas sarkasme dan non-sarkasme. Hasil ini mengindikasikan bahwa penggabungan representasi kontekstual IndoBERT dengan pengetahuan leksikal dari *Knowledge Graph* mampu meningkatkan kemampuan model dalam menangkap pola sarkasme implisit yang tidak dapat dideteksi hanya dari fitur linguistik permukaan.

3.2 Analisis Kurva Pelatihan

Gambar 3 memperlihatkan perubahan nilai training loss dan validation loss selama lima epoch

This Journal is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

ISSN: 3110-0864 (Media Online)
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | pp: 232-240
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.146

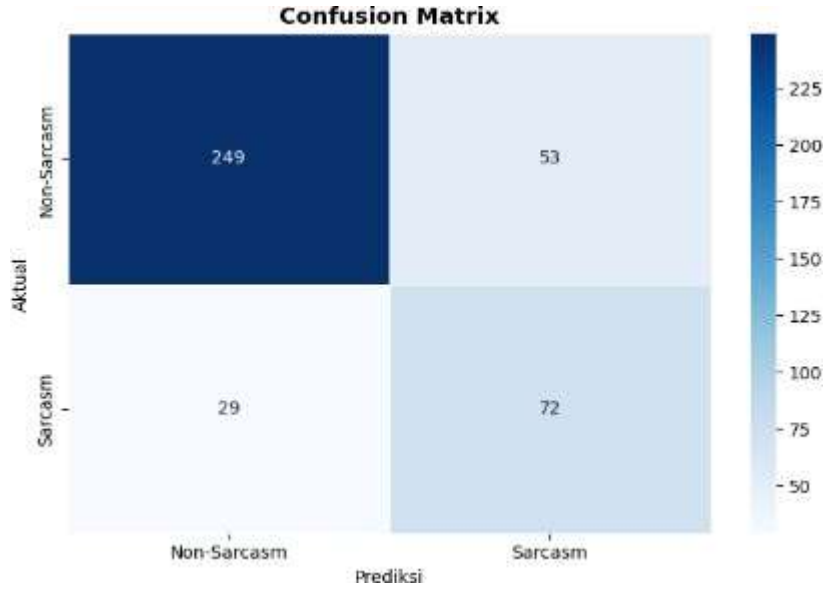


Gambar 2. Kurva Training History

Berdasarkan Gambar 3, nilai training loss mengalami penurunan secara bertahap dari epoch pertama hingga epoch terakhir. Validation loss juga menunjukkan kecenderungan yang stabil tanpa indikasi overfitting yang berarti. Hal ini menunjukkan bahwa kombinasi weight decay, dropout, dan linear warmup scheduler mampu membantu proses regularisasi model dengan baik. Nilai F1-score validasi terbaik diperoleh pada epoch ke-1 sebesar 84%.

3.3 Confusion Matrix

Gambar 4 menampilkan *confusion matrix* dari hasil prediksi model pada data uji.



Gambar 4. Confusion Matrix pada Data Uji

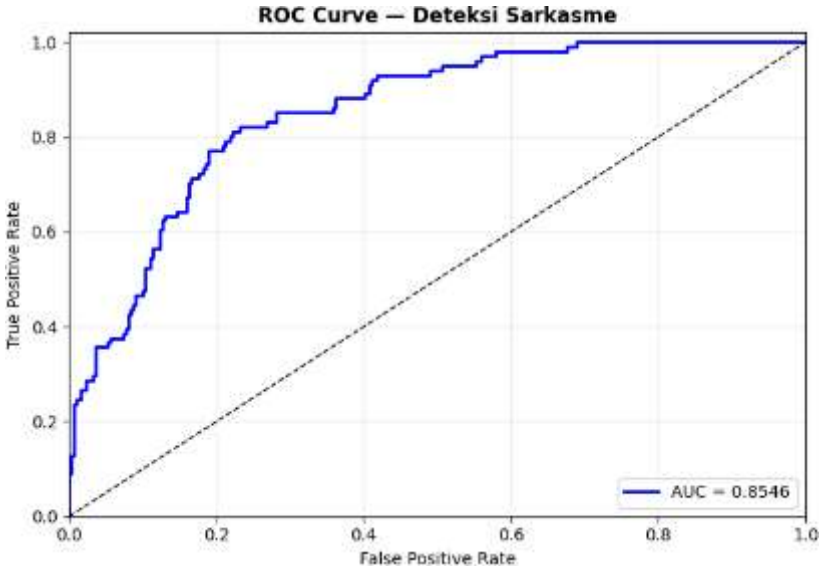
Berdasarkan Gambar 4, model berhasil mengklasifikasi 72 data sarkasme dengan benar *(true positive)* dan 249 data non-sarkasme dengan benar *(true negative).* Sebanyak 29 data termasuk dalam kategori *false negative,* yaitu data sarkasme yang yaitu data sarkasme yang diprediksi sebagai non-sarkasme. Kasus tersebut umumnya berupa sarkasme implisit yang tidak memiliki penanda kontradiksi yang jelas sehingga lebih sulit dikenali oleh model. Di sisi lain, terdapat 53 kasus *false positive*, yaitu data non-sarkasme yang diprediksi sebagai sarkasme karena mengandung ungkapan hiperbola atau kata-kata positif yang berlebihan.

This Journal is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

ISSN: 3110-0864 (Media Online)
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | pp: 232-240
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.146

3.4 Analisis Kurva ROC

Gambar 5 menunjukkan kurva ROC (*Receiver Operating Characteristic*) model pada data uji.



Gambar 5. Kurva ROC Model Hybrid IndoBERT-KG

Kurva ROC pada Gambar 5 menunjukkan nilai AUC (*Area Under Curve*) sebesar 85,46%. Nilai ini menunjukkan bahwa model memiliki kemampuan yang baik dalam memisahkan kelas sarkasme dari non-sarkasme pada berbagai ambang batas (*threshold*) klasifikasi. Kurva yang jauh dari garis diagonal *baseline* menandakan bahwa model jauh lebih baik dari klasifikasi acak.

3.5 Studi Kasus Prediksi

Untuk memberikan gambaran mengenai perilaku model secara kualitatif, Tabel 5 menyajikan beberapa contoh hasil prediksi beserta nilai confidence pada data uji.

Tabel 5. Contoh Hasil Prediksi Model

Teks	Label	Confidence
"Wah bagus banget pelayanannya, nunggunya cuma 3 jam!"	Sarcasm	97.72%
"Selamat ya udah berhasil ngacauin semua rencana"	Sarcasm	76.48%
"Film ini bagus banget, tidur saya jadi nyenyak"	Sarcasm	91.24%
"Makasih ya udah tepat waktu, cuma telat 2 jam aja"	Sarcasm	79.05%
"Cuaca hari ini cerah sekali, banjir di mana-mana"	Sarcasm	89.37%

Kalimat pertama pada Tabel 5, yaitu *"Wah bagus banget pelayanannya, nunggunya cuma 3 jam!"*, merupakan contoh sarkasme implisit yang mengandung pertentangan antara ungkapan positif *"bagus banget pelayanannya"* dan fakta negatif *"nunggunya cuma 3 jam"*. Model mampu mengenali pola tersebut karena Knowledge Graph menyimpan informasi kontradiksi semantik dan penanda sarkasme yang memperkuat sinyal sarkastik melalui mekanisme KG Attention.

Kalimat ketiga, *"Film ini bagus banget, tidur saya jadi nyenyak"*, menunjukkan bentuk sarkasme implisit yang lebih halus. Ungkapan positif *"bagus banget"* bertolak belakang dengan makna tersirat bahwa film tersebut membosankan hingga membuat penonton tertidur. Model berhasil mengenali pola tersebut melalui kombinasi representasi kontekstual IndoBERT dan informasi sentimen yang diperoleh dari Knowledge Graph.

Selain itu, komponen Implicit Sarcasm Detector pada arsitektur yang diusulkan turut berperan dalam mendeteksi sarkasme yang tidak memiliki penanda kontradiksi secara langsung pada Knowledge Graph, dengan memanfaatkan representasi yang telah diperkaya oleh mekanisme KG Attention.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem deteksi sarkasme implisit pada teks berbahasa Indonesia menggunakan arsitektur *Hybrid IndoBERT* dan *Knowledge Graph*. Model yang diusulkan mengintegrasikan representasi semantik kontekstual dari *IndoBERT* dengan pengetahuan leksikal domain-spesifik bahasa

Copyright © 2026 **Author**, Page 238
This Journal is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

ISSN: 3110-0864 (Media Online)
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | pp: 232-240
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.146

Indonesia yang direpresentasikan dalam *Knowledge Graph* melalui mekanisme *KG Attention*. *Knowledge Graph* yang dibangun mencakup leksikon sentimen, penanda sarkasme, pasangan kontradiksi semantik, dan kamus normalisasi *slang* bahasa Indonesia, yang secara bersama-sama menyediakan sinyal linguistik yang kaya untuk mendeteksi sarkasme implisit. Hasil eksperimen pada dataset *Reddit Indonesia Sarcastic* menunjukkan bahwa model yang diusulkan mencapai akurasi 79,65%, *F1-score* sebesar 80,31%, dan ROC-AUC sebesar 85,46%. Analisis kualitatif melalui studi kasus menunjukkan bahwa model mampu menangkap pola kontradiksi semantik yang menjadi ciri khas sarkasme implisit, termasuk kontradiksi antara ekspresi pujian dengan konteks situasi yang negatif. Mekanisme *KG Attention* terbukti efektif dalam memperkaya representasi kontekstual IndoBERT dengan pengetahuan leksikal, sehingga meningkatkan kemampuan model dalam membedakan sarkasme implisit dari teks non-sarkasme dengan kemiripan superfisial. Keterbatasan penelitian ini antara lain adalah *Knowledge Graph* yang dibangun secara manual sehingga cakupan leksikonnya terbatas, serta ketergantungan pada satu sumber dataset. Untuk penelitian selanjutnya, disarankan untuk memperluas *Knowledge Graph* menggunakan teknik *crowdsourcing* atau penambangan otomatis dari korpus bahasa Indonesia yang lebih besar. Selain itu, pengujian model pada domain yang lebih beragam seperti ulasan produk *e-commerce* dan komentar berita daring perlu dilakukan untuk mengukur kemampuan generalisasi model. Eksplorasi teknik augmentasi data juga disarankan untuk menangani potensi ketidakseimbangan kelas yang dapat mempengaruhi performa deteksi pada kelas minoritas.

REFERENCES

[1] E. Murakami, T. Shionoya, S. Komenoi, Y. Suzuki, and F. Sakane, “Cloning and characterization of novel testis-Specific diacylglycerol kinase η splice variants 3 and 4,” *PLoS One*, vol. 11, no. 9, Sep. 2016, doi: 10.1371/journal.pone.

[2] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, “IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP,” Online. [Online]. Available: https://huggingface.co/

[3] Z. Zhang, X. Han, Z. Liu, X. Jiang, M. Sun, and Q. Liu, “ERNIE: Enhanced Language Representation with Informative Entities.”

[4] R. A. Potamias, G. Siolas, and A. G. Stafylopatis, “A transformer-based approach to irony and sarcasm detection,” *Neural Comput. Appl.*, vol. 32, no. 23, pp. 17309–17320, Dec. 2020, doi: 10.1007/s00521-020-05102-3.

[5] E. Wallace, M. Gardner, and S. Singh, “Interpreting Predictions of NLP Models,” in *EMNLP 2020 - Conference on Empirical Methods in Natural Language Processing, Tutorial Abstracts*, Association for Computational Linguistics (ACL), 2020, pp. 20–23. doi: 10.18653/v1/P17.

[6] M. A. A. Saputra, A. Alamsyah, D. P. Ramadhani, T. S. Siadari, and H. Fakhrurroja, “IndoBERT-Sentiment: Context-Conditioned Sentiment Classification for Indonesian Text,” Apr. 2026, [Online]. Available: http://arxiv.org/abs/2604.07057

[7] B. Wilie *et al.*, “IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding.” [Online]. Available: https://github.com/annisanurulazhar/absa-playground

[8] W. W. and A. T. H. D. Suhartono, “Hugging Face.” Accessed: Jun. 14, 2026. [Online]. Available: w11wo/reddit_indonesia_sarcastic

[9] S. K. Emmanuel Daniel Widhiarto, “Knowledge Graph Struktur Data yang Membuat Mesin Paham Dunia.” Accessed: Jun. 14, 2026. [Online]. Available: https://socs.binus.ac.id/2025/11/26/knowledge-graph-struktur-data-yang-membuat-mesin-paham-dunia/

[10] M. E. Peters *et al.*, “Knowledge Enhanced Contextual Word Representations.” Accessed: Jun. 14, 2026. [Online]. Available: https://medium.com/swlh/recall-precision-f1-roc-auc-and-everything-542aedf322b9

[11] P. Luo, X. Zhu, T. Xu, Y. Zheng, and E. Chen, “Semantic Interaction Matching Network for Few-Shot Knowledge Graph Completion,” *ACM Transactions on the Web*, vol. 18, no. 2, Jan. 2024, doi: 10.1145/3589557.

This Journal is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

ISSN: 3110-0864 (Media Online)
JURNAL ILMU KOMPUTER DAN TEKNIK INFROMATIKA
Volume 2 No 2 Juli 2026 | pp: 232-240
https://journals.raskhamedia.or.id/index.php/juikti
DOI: https://doi.org/10.64803/juikti.v2i2.146

[12] Liu Ye, Wan Yao, He Lifang, Peng Hao, and Yu Philip S., "KG-BART: Knowledge Graph-Augmented BART for Generative Commonsense Reasoning." Accessed: Jun. 14, 2026. [Online]. Available: https://huggingface.co/papers/2009.12677

[13] I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization," Jan. 2019, [Online]. Available: http://arxiv.org/abs/1711.05101

[14] Shalev Ofir, "Recall, Precision, F1, ROC, AUC, and everything."

[15] W. Liu *et al.*, "The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20) K-BERT: Enabling Language Representation with Knowledge Graph." [Online]. Available: www.aaai.org

Copyright © 2026 **Author**, Page 240
This Journal is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.