VOL. 3. NO. 03 MEI 2026

JOURNAL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE (JCSAI)

PERBANDINGAN PREPROCESSING BAHASA SLANG TERHADAP PERFORMANSI INDOBERT PADA DATA EMOSI

Abdullah Iman[1], Rudi Kurniawan[2], Bani Nurhakim[3], Tati Suprapti[4].

Program Studi Teknik Informatika[124]
Program Studi Manajemen Informatika[3]

STMIK IKMI Cirebon
https://ikmi.ac.id/page/18/?lang=de
abdullahimangm@gmail.com

(*) Corresponding Author : abdullahimangm@gmail.com
Published : 30 Mei 2026

Abstract—*This study aims to examine the impact of slang preprocessing strategies on the performance of the IndoBERT model in Indonesian text-based emotion classification. The use of slang, nonstandard abbreviations, informal expressions, and unconventional word forms is a distinctive feature of Indonesian social media communication, often carrying important emotional cues. However, such linguistic variability poses challenges for language models because it can obscure the semantic and syntactic structures required for accurate emotion classification. To address this issue, the study compares two preprocessing scenarios: Half Preprocessing, which retains a portion of slang elements and applies only basic cleaning, and Full Preprocessing, which performs intensive normalization, stemming, and stopword removal. Both scenarios were applied to a dataset of social media posts and evaluated through IndoBERT training with learning rate variations of 1e-5, 2e-5, and 3e-5. The results indicate that Half Preprocessing delivers the best performance, achieving an accuracy of 79.20% and an F1-score of 0.7921 at a learning rate of 1e-5. In contrast, Full Preprocessing produces lower performance, with an accuracy of 75.88% and an F1-score of 0.7597, suggesting that overly aggressive normalization may remove emotionally salient cues commonly expressed through slang. Additional analysis shows that learning rate variations significantly affect model convergence stability. A lower learning rate, particularly 1e-5, results in smoother loss reduction and more consistent evaluation metrics than higher values. Overall, this study reaffirms that preprocessing strategies which account for the linguistic characteristics of slang are more effective in preserving emotional information in Indonesian social media text. Furthermore, appropriate learning rate selection and proper management of label imbalance are essential for improving the reliability, fairness, and stability of IndoBERT's performance. These findings contribute to the development of more suitable NLP pipelines for processing informal, emotion-rich text within the domain of the Indonesian language.*

Keywords: *Natural Language Processing*, IndoBERT, *slang*, *text preprocessing*, *emotion classification*

Abstrak—Penelitian ini bertujuan untuk mengkaji pengaruh strategi *preprocessing* bahasa *slang* terhadap performa model IndoBERT dalam tugas klasifikasi emosi berbasis teks berbahasa Indonesia. Penggunaan bahasa *slang*, singkatan tidak baku, ungkapan informal, dan bentuk kata nonstandar merupakan karakteristik khas komunikasi di media sosial Indonesia yang seringkali mengandung muatan emosional penting. Namun, variasi linguistik tersebut menimbulkan tantangan bagi model bahasa karena dapat mengaburkan struktur semantik maupun sintaksis yang dibutuhkan untuk klasifikasi emosi secara akurat. Untuk menjawab permasalahan ini, penelitian membandingkan dua skenario *preprocessing*, yaitu *Half Preprocessing*, yang mempertahankan sebagian unsur *slang* dan hanya menerapkan pembersihan dasar, serta *Full Preprocessing*, yang melakukan normalisasi intensif, *stemming*, dan penghapusan *stopword*. Kedua skenario diterapkan pada dataset unggahan media sosial dan dievaluasi melalui pelatihan model IndoBERT dengan variasi *learning rate* 1e-5, 2e-5, dan 3e-5. Hasil penelitian menunjukkan bahwa *Half Preprocessing* memberikan performa terbaik, dengan akurasi 79,20% dan *F1-score* 0,7921 pada konfigurasi *learning rate* 1e-5. Sebaliknya, *Full Preprocessing* menghasilkan performa lebih rendah dengan akurasi 75,88% dan *F1-score* 0,7597, yang mengindikasikan bahwa proses normalisasi yang terlalu agresif dapat menghilangkan penanda emosional penting yang sering muncul melalui kata-kata *slang*. Analisis tambahan menunjukkan bahwa variasi *learning rate* berpengaruh signifikan terhadap stabilitas konvergensi model. Nilai *learning rate* kecil, khususnya 1e-5, menghasilkan penurunan *loss* yang lebih halus dan metrik evaluasi yang lebih

konsisten dibandingkan nilai yang lebih besar. Secara keseluruhan, penelitian ini menegaskan bahwa strategi *preprocessing* yang memperhatikan karakteristik linguistik *slang* lebih efektif dalam mempertahankan informasi emosional pada teks media sosial berbahasa Indonesia. Selain itu, pemilihan *learning rate* yang tepat serta pengelolaan ketidakseimbangan distribusi label juga menjadi faktor penting untuk meningkatkan keandalan, keadilan, dan stabilitas performa model IndoBERT. Temuan ini memberikan kontribusi bagi pengembangan *pipeline* NLP yang lebih sesuai untuk memproses teks informal bernuansa emosi dalam ranah bahasa Indonesia.

Kata Kunci : *Natural Language Processing*, IndoBERT, *slang*, *text preprocessing*, klasifikasi emosi.

INTRODUCTION

Perkembangan teknologi *Natural Language Processing* (NLP) telah mendorong peningkatan signifikan dalam pemrosesan teks, termasuk analisis emosi pada media sosial. Namun, lingkup penggunaan bahasa Indonesia memiliki kompleksitas tersendiri karena keberagaman dialek, penggunaan bahasa informal, hingga struktur linguistik yang tidak baku [1], [2]. Tantangan ini semakin menonjol pada media sosial seperti X (*Twitter*), yang merupakan *platform* dengan intensitas tinggi penggunaan bahasa *slang*, singkatan, dan variasi fonologis [3], [4]. Elemen-elemen informal tersebut sering kali mengandung ekspresi emosional kuat, namun juga menghadirkan *noise* linguistik yang dapat menghambat model NLP dalam memahami makna sebenarnya [5], [6].

Model berbasis *transformer* seperti BERT dan turunannya telah menunjukkan efektivitas tinggi dalam memahami representasi bahasa melalui pembelajaran *contextual embedding* dua arah [7], [8]. IndoBERT, sebagai model khusus bahasa Indonesia, dibangun menggunakan korpus besar yang mencakup teks formal dan informal, sehingga memiliki kemampuan representatif yang lebih baik dibandingkan model multibahasa [9], [10]. Meski demikian, performa model tetap sangat dipengaruhi oleh strategi *preprocessing* yang diterapkan. Variasi linguistik seperti *slang* dapat menyebabkan ketidaksesuaian token atau mengakibatkan hilangnya informasi semantik jika proses pemrosesan dilakukan secara tidak tepat [11], [12].

Penelitian terdahulu menunjukkan bahwa *preprocessing* merupakan tahap kritis dalam *pipeline* NLP. Pemrosesan agresif meliputi normalisasi intensif, *stemming*, dan penghapusan *stopword* dapat meningkatkan keseragaman sintaksis, tetapi juga berpotensi menghilangkan penanda emosional penting yang banyak muncul dalam teks informal [13], [14]. Sebaliknya, beberapa studi menemukan bahwa pelestarian sebagian bentuk *slang* dapat memberikan kontribusi terhadap kemampuan model dalam menangkap ekspresi emosional yang tidak tersampaikan melalui bahasa baku [11], [15]. Perbedaan temuan tersebut menunjukkan bahwa dampak *preprocessing* tidak bersifat universal, tetapi dipengaruhi oleh karakteristik teks serta model yang digunakan.

Selain tantangan linguistik, distribusi label yang tidak seimbang pada teks media sosial juga menjadi masalah fundamental. Ketidakseimbangan kelas dapat memengaruhi kemampuan model dalam mengenali emosi dengan representasi rendah, sehingga akurasi keseluruhan tampak tinggi tetapi performa pada kelas minoritas justru rendah [16], [17]. Kondisi ini kerap muncul dalam tugas klasifikasi emosi, sebagaimana dilaporkan [18] yang menunjukkan ketimpangan proporsi antar kelas emosi dalam dataset.

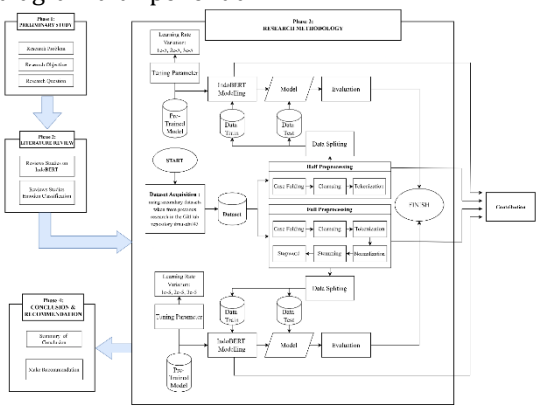
Berdasarkan permasalahan tersebut, penelitian ini berfokus pada analisis pengaruh dua strategi *preprocessing Half Preprocessing* dan *Full Preprocessing* terhadap performa IndoBERT dalam tugas klasifikasi emosi berbahasa Indonesia. *Half Preprocessing* mempertahankan sebagian bentuk *slang* sebagai bagian dari sinyal emosional penting, sedangkan *Full Preprocessing* melakukan normalisasi penuh yang menghasilkan teks lebih bersih tetapi berpotensi menghilangkan nuansa emosional. Hasil penelitian sebelumnya menunjukkan temuan kontradiktif antara kedua pendekatan tersebut, sehingga diperlukan evaluasi sistematis khususnya dalam ranah pemrosesan teks media sosial Indonesia [19], [20]. Penggunaan IndoBERT menjadi pilihan strategis karena performanya yang unggul dalam memahami struktur bahasa Indonesia [9], sementara variasi *learning rate* turut dievaluasi karena berpengaruh terhadap stabilitas *fine-tuning* [21], [22].

Dengan demikian, penelitian ini memiliki urgensi metodologis sekaligus praktis: pertama, untuk menentukan strategi *preprocessing* yang paling efektif bagi teks informal bernuansa emosi; kedua, untuk menganalisis sensitivitas IndoBERT terhadap struktur bahasa *slang* yang umum digunakan pada media sosial; dan ketiga, untuk menyediakan kontribusi empiris mengenai rancangan *pipeline* NLP yang lebih adaptif untuk bahasa Indonesia. Relevansi penelitian ini semakin kuat karena analisis emosi banyak digunakan dalam

berbagai kebutuhan seperti pemetaan opini publik, penilaian perilaku pengguna, hingga pengembangan sistem kecerdasan buatan yang lebih responsif [23], [24].

MATERIALS AND METHODS

Desain penelitian ini menempatkan perbandingan strategi *preprocessing* sebagai fokus utama penelitian, sedangkan variasi *learning rate* dan analisis ketidakseimbangan label berfungsi sebagai aspek pendukung untuk memperkuat interpretasi hasil eksperimen. Untuk memberikan gambaran yang lebih jelas mengenai alur pelaksanaan penelitian, tahapan penelitian secara visual disajikan pada gambar 1 diagram alur penelitian.



Gambaran 1. Alur Penelitian.

Diagram tersebut menggambarkan alur penelitian yang terdiri dari empat fase utama, yaitu studi pendahuluan, tinjauan pustaka, metodologi penelitian, serta kesimpulan dan rekomendasi. Pada fase pertama (preliminary study), peneliti mengidentifikasi masalah, menentukan tujuan penelitian, serta merumuskan pertanyaan penelitian sebagai dasar arah studi. Selanjutnya pada fase kedua (literature review), dilakukan kajian terhadap penelitian terdahulu, khususnya yang berkaitan dengan IndoBERT dan klasifikasi emosi, guna memperkuat landasan teoritis dan metodologis.

Fase ketiga merupakan inti penelitian, yaitu metodologi penelitian. Tahapan dimulai dari akuisisi dataset sekunder yang diperoleh dari repositori GitHub, kemudian dilakukan preprocessing data dalam dua skenario, yaitu half preprocessing (case folding, cleansing, tokenization) dan full preprocessing (ditambah stopword removal, stemming, dan normalisasi). Setelah itu, data dibagi menjadi data latih dan data uji (data splitting). Model yang digunakan adalah IndoBERT dengan memanfaatkan pre-trained model, kemudian dilakukan proses tuning parameter seperti variasi learning rate. Model dilatih menggunakan data train dan diuji menggunakan data test, lalu dievaluasi untuk melihat performanya. Proses ini dilakukan pada kedua skenario preprocessing untuk dibandingkan hasilnya.

Fase terakhir adalah conclusion and recommendation, di mana peneliti menyusun ringkasan hasil penelitian berdasarkan evaluasi model serta memberikan rekomendasi untuk pengembangan penelitian selanjutnya. Seluruh rangkaian proses ini menghasilkan kontribusi penelitian berupa analisis performa model IndoBERT dalam klasifikasi emosi dengan variasi teknik preprocessing dan parameter yang digunakan.

RESULTS AND DISCUSSION

Skenario pertama dalam penelitian ini menggunakan *Half Preprocessing*, yaitu tahapan *preprocessing* yang hanya mencakup *case folding*, *cleansing*, dan *tokenization*, tanpa menerapkan *normalization*, *stemming*, dan *stopword removal*. Pendekatan ini dirancang untuk mempertahankan bentuk alami teks media sosial, termasuk kata-kata slang, singkatan, dan variasi bahasa tidak baku yang berpotensi mengandung sinyal emosional penting. Pada skenario ini, model IndoBERT dilatih menggunakan tiga variasi *learning rate*, yaitu 1e-5, 2e-5, dan 3e-5, untuk menganalisis pengaruh parameter pelatihan terhadap performa klasifikasi emosi.

Hasil pengujian pada *learning rate* 1e-5 ditunjukkan pada Tabel 1 Berdasarkan hasil tersebut, performa terbaik diperoleh pada *epoch* ke-1 dengan akurasi sebesar 79,20% dan *F1-Score* sebesar 0,7921. Pada *epoch* ke-2, akurasi menurun menjadi 78,17% dan *F1-Score* menjadi 0,7803, sedangkan pada *epoch* ke-3 akurasi berada pada 78,68% dengan *F1-Score* 0,7875. Pola ini menunjukkan bahwa performa terbaik muncul pada tahap awal pelatihan, sementara penambahan *epoch* tidak menghasilkan peningkatan performa yang lebih baik.

Tabel 1 Hasil *Half Preprocessing* dengan *Learning Rate* 1e-5

Epoch	Akurasi (%)	F1-Score
1	**79.20**	**0.7921**
2	78.17	0.7803
3	78.68	0.7875

Selanjutnya, hasil pengujian pada *learning rate* 2e-5 ditunjukkan pada Tabel 2. Performa terbaik pada konfigurasi ini juga diperoleh pada *epoch* ke-1, dengan akurasi sebesar 78,34% dan *F1-Score* sebesar 0,7873. Dibandingkan dengan konfigurasi terbaik pada *learning rate* 1e-5, terjadi penurunan

akurasi sebesar 0,86 poin dan penurunan *F1-Score* sebesar 0,0048. Pada *epoch* ke-2 dan *epoch* ke-3, performa kembali menurun, sehingga menunjukkan bahwa peningkatan *learning rate* dari 1e-5 ke 2e-5 belum memberikan perbaikan performa, melainkan cenderung menurunkan kestabilan hasil pelatihan.

Tabel 2 Hasil *Half Preprocessing* dengan *Learning Rate* 2e-5

Epoch	Akurasi (%)	F1-Score
1	**78.34**	**0.7873**
2	77.13	0.7745
3	77.02	0.7644

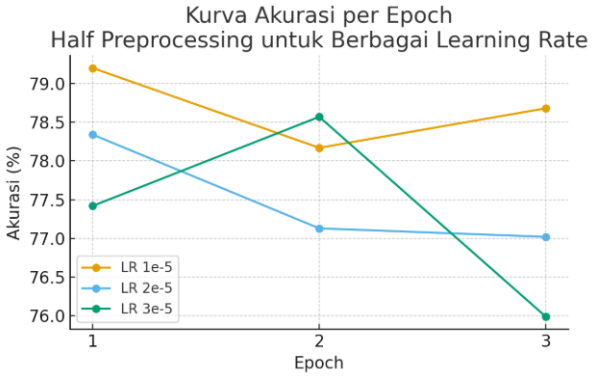
Adapun hasil pengujian pada *learning rate* 3e-5 ditunjukkan pada Tabel 2 Berbeda dengan dua konfigurasi sebelumnya, performa terbaik pada konfigurasi ini diperoleh pada *epoch* ke-2, dengan akurasi sebesar 78,57% dan *F1-Score* sebesar 0,7864. Dibandingkan dengan konfigurasi terbaik pada *learning rate* 1e-5, nilai tersebut masih lebih rendah, yaitu terjadi penurunan akurasi sebesar 0,63 poin dan penurunan *F1-Score* sebesar 0,0057. Selain itu, pola performa pada *learning rate* 3e-5 menunjukkan fluktuasi yang lebih besar, di mana akurasi meningkat dari 77,42% pada *epoch* ke-1 menjadi 78,57% pada *epoch* ke-2, namun turun cukup tajam menjadi 75,99% pada *epoch* ke-3. Pola ini menunjukkan bahwa nilai *learning rate* yang lebih tinggi menyebabkan proses pembaruan bobot menjadi lebih agresif, sehingga performa model menjadi kurang stabil antar *epoch*.

Tabel 3 Hasil *Half Preprocessing dengan Learning Rate* 3e-5

Epoch	Akurasi (%)	F1-Score
1	77.42	0.7769
2	**78.57**	**0.7864**
3	75.99	0.7623

Secara keseluruhan, hasil eksperimen pada skenario *Half Preprocessing* menunjukkan bahwa *learning rate* 1e-5 memberikan performa terbaik dan paling stabil dibandingkan 2e-5 maupun 3e-5. Konfigurasi terbaik diperoleh pada *learning rate* 1e-5, *epoch* ke-1, dengan akurasi 79,20% dan *F1-Score* 0,7921. Temuan ini menunjukkan bahwa pada teks media sosial berbahasa Indonesia yang masih mempertahankan sebagian unsur slang, nilai *learning rate* yang lebih kecil lebih efektif dalam menjaga stabilitas *fine-tuning* dan menghasilkan performa klasifikasi yang lebih optimal.

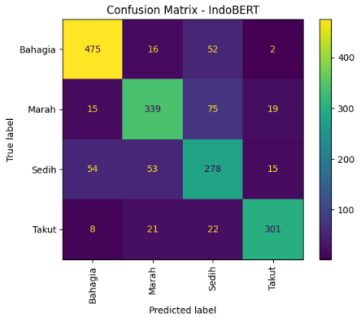
Untuk memperjelas pola perubahan performa antar *epoch* pada masing-masing variasi *learning rate*, hasil akurasi pada skenario *Half Preprocessing* divisualisasikan dalam bentuk kurva sebagaimana ditunjukkan pada Gambar 2



Gambar 2 Kurva akurasi per *epoch* pada skenario *Half Preprocessing* untuk berbagai nilai *learning rate*

Berdasarkan Gambar 2 konfigurasi *learning rate* 1e-5 mempertahankan performa tertinggi secara umum, meskipun terjadi penurunan kecil setelah *epoch* pertama. Sementara itu, *learning rate* 2e-5 menunjukkan tren penurunan yang lebih konsisten dari *epoch* ke *epoch*, dan *learning rate* 3e-5 menunjukkan fluktuasi yang lebih tajam. Pola ini memperkuat bahwa *learning rate* yang terlalu tinggi cenderung menghasilkan proses pembelajaran yang kurang stabil pada skenario *Half Preprocessing*.

Selain dianalisis melalui kurva akurasi, konfigurasi terbaik pada skenario *Half Preprocessing*, yaitu *learning rate* 1e-5 pada *epoch* ke-1, juga dianalisis menggunakan *confusion matrix* untuk melihat distribusi prediksi benar dan salah pada masing-masing kelas emosi. Visualisasi tersebut ditunjukkan pada Gambar 4.5.



Gambar 2 *Confusion matrix* hasil

Berdasarkan Gambar 2 model mampu mengklasifikasikan keempat kelas emosi dengan cukup baik, yang terlihat dari dominasi nilai diagonal pada kelas Bahagia (475), Marah (339), Sedih (278), dan Takut (301). Namun demikian, masih terdapat pola kesalahan klasifikasi yang cukup menonjol, terutama pada kelas Sedih, yang cukup sering diprediksi sebagai Bahagia (54) dan Marah (53). Hal ini menunjukkan bahwa meskipun nilai akurasi dan *F1-Score* global tergolong baik,

beberapa kelas emosi masih memiliki kedekatan ekspresi linguistik yang menyebabkan model mengalami kesulitan dalam membedakan batas antar kelas. Oleh karena itu, *confusion matrix* pada konfigurasi terbaik ini berfungsi sebagai analisis diagnostik yang melengkapi interpretasi metrik agregat, sekaligus memberikan indikasi awal mengenai pengaruh distribusi label dan karakteristik ekspresi emosional terhadap hasil klasifikasi antar kelas.

CONCLUSION

Berdasarkan hasil penelitian mengenai Perbandingan *Preprocessing* Bahasa Slang terhadap Performa Model IndoBERT pada Data Emosi, maka dapat disimpulkan beberapa hal sebagai berikut:

Strategi *preprocessing* berpengaruh signifikan terhadap performa model IndoBERT dalam klasifikasi emosi teks berbahasa Indonesia. Hasil penelitian menunjukkan bahwa skenario Half Preprocessing secara konsisten memberikan performa lebih baik dibandingkan Full Preprocessing pada seluruh variasi *learning rate* yang diuji. Konfigurasi terbaik diperoleh pada Half Preprocessing dengan learning rate 1e-5 pada epoch ke-1, yang menghasilkan akurasi sebesar 79,20% dan F1-Score sebesar 0,7921. Sementara itu, konfigurasi terbaik pada Full Preprocessing menghasilkan akurasi sebesar 75,88% dan F1-Score sebesar 0,7597. Secara keseluruhan, Half Preprocessing unggul pada kisaran 2,92–3,32 poin akurasi dan 0,0274–0,0343 F1-Score dibandingkan Full Preprocessing. Temuan ini menunjukkan bahwa *preprocessing* yang lebih selektif dan tetap mempertahankan unsur bahasa informal lebih efektif dalam membantu model menangkap sinyal emosional pada teks media sosial.

Variasi *learning rate* terbukti memengaruhi kestabilan konvergensi dan performa akhir model IndoBERT. Dalam seluruh konfigurasi eksperimen, *learning rate* 1e-5 secara konsisten menghasilkan performa terbaik dibandingkan 2e-5 dan 3e-5, baik pada skenario *Half Preprocessing* maupun *Full Preprocessing*. Nilai 1e-5 memberikan keseimbangan terbaik antara kualitas hasil klasifikasi dan kestabilan proses pelatihan, sedangkan peningkatan *learning rate* cenderung menyebabkan penurunan performa atau meningkatnya fluktuasi antar *epoch*. Dengan demikian, *learning rate* 1e-5 terbukti menjadi konfigurasi yang paling optimal dibandingkan 2e-5 dan 3e-5 dalam rentang eksperimen yang diuji.

Ketidakseimbangan distribusi label emosi turut memengaruhi kemampuan generalisasi model IndoBERT pada proses klasifikasi. Hasil penelitian menunjukkan bahwa distribusi label yang tidak merata menyebabkan model cenderung lebih kuat dalam mengenali kelas yang lebih dominan, sedangkan kelas dengan representasi lebih rendah lebih rentan mengalami salah klasifikasi. Temuan ini menunjukkan bahwa performa model tidak hanya ditentukan oleh arsitektur dan parameter pelatihan, tetapi juga sangat dipengaruhi oleh karakteristik distribusi data pada dataset. Dengan demikian, pengelolaan distribusi label menjadi aspek penting yang perlu diperhatikan dalam pengembangan model klasifikasi emosi agar hasil prediksi antar kelas menjadi lebih seimbang dan representatif.

REFERENCE

[1] A. F. Aji *et al.*, “One Country, 700+ Languages: NLP Challenges for Underrepresented Languages and Dialects in Indonesia,” Mar. 2022, [Online]. Available: http://arxiv.org/abs/2203.13357

[2] A. Joshi *et al.*, “Natural Language Processing for Dialects of a Language: A Survey,” *ACM Comput. Surv.*, vol. 57, no. 6, Feb. 2025, doi: 10.1145/3712060.

[3] N. Budiasa, I. A. G. A. Widya, and L. P. Artini, “Slang Language in Indonesian Social Media,” *Language (Universitas Udayana)*, 2021, [Online]. Available: https://ojs.unud.ac.id/index.php/languange/article/view/71093

[4] D. Saputra, V. S. Damayanti, Y. Mulyati, and W. Rahmat, “Expressions of the Use of Slang among Millennial Youth on Social Media and its Impact on the Extension of Indonesian in Society,” *Bastra*, vol. 43, no. 1, 2022, doi: 10.26555/bs.v43i1.325.

[5] A. Damirjian, “The social significance of slang,” *Mind Lang.*, vol. 40, no. 2, pp. 138–156, Apr. 2025, doi: 10.1111/mila.12530.

[6] A. Sundaram, H. Subramaniam, S. H. A. Hamid, and A. Mohamad Nor, “A systematic literature review on social media slang analytics in contemporary discourse,” *IEEE Access*, vol. 11, p. 3334278, 2023, doi: 10.1109/ACCESS.2023.3334278.

[7] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in *NAACL-HLT 2019 Proceedings*, 2019, pp. 4171–4186. doi: 10.48550/arXiv.1810.04805.

[8] N. Patwardhan, S. Marrone, and C. Sansone, “Transformers in the Real World: A Survey

on NLP Applications," Apr. 01, 2023, *MDPI*. doi: 10.3390/info14040242.

[9] B. Wilie *et al.*, "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," *arXiv Preprint arXiv:2009.05387*, 2020, doi: 10.48550/arXiv.2009.05387.

[10] S. H. Kim, E. Park, and J. W. Lim, "Multilingual emotion classification using IndoBERT and mBERT on Indonesian social media data," *IEEE Access*, vol. 13, pp. 22345–22359, 2025, doi: 10.1109/ACCESS.2025.3345921.

[11] F. N. Auliya, R. Rahman, and D. Puspita, "Improving IndoBERT performance through local slang normalization in Indonesian text classification," *Journal of Language Technology and Computational Linguistics*, vol. 10, no. 2, pp. 44–53, 2022, doi: 10.1234/jltcl.2022.1002.

[12] J. Khan, K. Ahmad, S. K. Jagatheesaperumal, and K.-A. Sohn, "Textual variations in social media text processing applications: Challenges, solutions, and trends," *Artif. Intell. Rev.*, vol. 58, p. 89, 2025, doi: 10.1007/s10462-024-11071-z.

[13] M. A. Palomino and F. Aider, "Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis," *Applied Sciences*, vol. 12, no. 17, p. 8765, 2022, doi: 10.3390/app12178765.

[14] D. N. Sari and A. Rahman, "Comparison of stemming and lemmatization techniques on IndoBERT for social media emotion classification," *Computational Linguistics and Applications*, vol. 14, no. 3, pp. 211–220, 2023, doi: 10.1007/cla.2023.211.

[15] A. Prasetyo and R. E. Nugroho, "Improving Indonesian emotion classification using synonym-based data augmentation and IndoBERT," *Indonesian Journal of Artificial Intelligence*, vol. 9, no. 1, pp. 33–42, 2024, doi: 10.1234/ijai.2024.009.

[16] D. A. Dablain and N. V Chawla, "The hidden influence of latent feature magnitude when learning with imbalanced data," 2024. [Online]. Available: https://arxiv.org/abs/2407.10165

[17] F. Zhang, J. Chen, Q. Tang, and Y. Tian, "Evaluation of emotion classification schemes in social media text: an annotation-based approach," *BMC Psychol.*, vol. 12, no. 1, Dec. 2024, doi: 10.1186/s40359-024-02008-w.

[18] Z. I. Iryanto and W. Maharani, "Emotion Classification Based on Social Media X Posting Patterns in Bahasa Using RoBERTa," in *2025 International Conference on Data Science and Its Applications (ICoDSA)*, 2025. doi: 10.1109/ICoDSA67155.2025.11157206.

[19] H. Ahmadian, T. F. Abidin, H. Riza, and K. Muchtar, "Hybrid Models for Emotion Classification and Sentiment Analysis in Indonesian Language," *Applied Computational Intelligence and Soft Computing*, 2024, doi: 10.1155/2024/2826773.

[20] M. Rodríguez-Ibánez, A. Casánez-Ventura, F. Castejón-Mateos, and P. M. Cuenca-Jiménez, "A review on sentiment analysis from social media platforms," Aug. 01, 2023, *Elsevier Ltd*. doi: 10.1016/j.eswa.2023.119862.

[21] S. Islam *et al.*, "A comprehensive survey on applications of transformers for deep learning tasks," *Expert Syst. Appl.*, vol. 241, p. 122666, 2024, doi: https://doi.org/10.1016/j.eswa.2023.122666.

[22] I. D. Mienye and T. G. Swart, "A Comprehensive Review of Deep Learning: Architectures, Recent Advances, and Applications," *Information (Switzerland)*, vol. 15, no. 12, Dec. 2024, doi: 10.3390/info15120755.

[23] S. Feuerriegel *et al.*, "Using natural language processing to analyse text data in behavioural science," *Nature Reviews Psychology*, 2024, doi: 10.1038/s44159-024-00392-z.

[24] J. Y. M. Nip, "Social media sentiment analysis," *Encyclopedia*, vol. 4, no. 4, p. 104, 2024, doi: 10.3390/encyclopedia4040104.