

DAFTAR PUSTAKA

- Aditya, B. R. (2015). *Penggunaan Web Crawler Untuk Menghimpun Tweets dengan Metode Pre-Processing Text Mining*. 7(2), 8.
- Adiwidya, B. M. D. (2009). *Algoritma Levenshtein dalam Pendekatan Approximate String Matching*. 4.
- Akbarianto Wibowo, H., Aziz Prawiro, T., Ihsan, M., Fikri Aji, A., Eko Prasajo, R., Mahendra, R., & Fitriany, S. (2020). Semi-Supervised Low-Resource Style Transfer of Indonesian Informal to Formal Language with Iterative Forward-Translation. Dalam *ArXiv e-prints*.
<https://ui.adsabs.harvard.edu/abs/2020arXiv201103286A>
- Atmaja, R. M. R. W. P. K., & Yustanti, Y. (2021). Analisis Sentimen Customer Review Aplikasi Ruang Guru dengan Metode BERT (Bidirectional Encoder Representations from Transformers). *Journal of Emerging Information Systems and Business Intelligence*, 02(03), 55–62.
- Buntoro, G. A. (2017). Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter. *Integer Journal*, 2, 32–41.
- Fimoza, D. (2021). *ANALISIS SENTIMEN TERHADAP FILM INDONESIA DENGAN PENDEKATAN BERT*. 95.
- Gligorić, K., Anderson, A., & West, R. (2020). Adoption of Twitter's New Length Limit: Is 280 the New 140? *ArXiv:2009.07661 [Cs]*.
<http://arxiv.org/abs/2009.07661>
- Hidayatullah, A. F. (2015). Language tweet characteristics of Indonesian citizens. *2015 International Conference on Science and Technology (TICST)*, 397–401. <https://doi.org/10.1109/TICST.2015.7369393>

- Khomsah, S. & Agus Sasmito Aribowo. (2020). Text-Preprocessing Model Youtube Comments in Indonesian. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 4(4), 648–654.
<https://doi.org/10.29207/resti.v4i4.2035>
- Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP. *ArXiv:2011.00677 [Cs]*. <http://arxiv.org/abs/2011.00677>
- Lourentzou, I., Manghnani, K., & Zhai, C. (2019). *Adapting Sequence to Sequence Models for Text Normalization in Social Media*. 11.
- Murthy, D. (2012). Towards a Sociological Understanding of Social Media: Theorizing Twitter. *Sociology*, 46(6), 1059–1073.
<https://doi.org/10.1177/0038038511422553>
- Nugraha, I. G. B. B., & Rizqullah, R. D. (2019). Normalisasi Kata Tidak Baku yang Tidak Disingkat dengan Jarak Perubahan. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, 8(3), 218.
<https://doi.org/10.22146/jnteti.v8i3.516>
- Nur, M. A. (2021). Perbandingan Levenshtein Distance Dan Jaro-Winkler Distance Untuk Koreksi Kata Dalam Preprocessing Analisis Sentimen Pengguna Twitter. *Jurnal Fokus Elektroda : Energi Listrik, Telekomunikasi, Komputer, Elektronika dan Kendali*, 6(2), 88.
<https://doi.org/10.33772/jfe.v6i2.17751>
- Rosmala, D., & Risyad, Z. M. (2017). Algoritma Levenshtein Distance dalam Aplikasi Pencarian Kata Isu di Kota Bandung pada Twitter. *MIND Journal*, 2(2), 1–2. <https://doi.org/10.26760/mindjournal>

- Samudro, A. A. (2019). *Normalisasi Teks Bahasa Indonesia pada Media Sosial Berdasarkan FastText Embeddings* [Undergraduate, Institut Teknologi Sepuluh Nopember]. <https://repository.its.ac.id/60580/>
- Sejarah dan Profil Perusahaan SiCepat Ekspres* | *Media62.ID*. (2021, September 14). <https://media62.id/karir/sicepat-ekspres/>
- Twitter: Most users by country*. (2022, Januari). Statista. <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>
- Wibowo, H. A., Nityasya, M. N., Akyürek, A. F., Fitriany, S., Aji, A. F., Prasajo, R. E., & Wijaya, D. T. (2021). IndoCollex: A Testbed for Morphological Transformation of Indonesian Word Colloquialism. *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 3170–3183. <https://doi.org/10.18653/v1/2021.findings-acl.280>

LAMPIRAN

1. *Source code*

Source code dan *dataset* yang digunakan pada penelitian ini dapat dilihat pada tautan berikut: <https://github.com/dikisiswanto/nlp-skripsi>

2. Kamus IndoCollex

Kamus kata slang yang digunakan pada penelitian ini bersumber dari *repository* IndoCollex yang dapat diakses pada tautan: <https://github.com/haryoa/indo-collex>

3. Kamus kata dasar KBBI

Kamus kata dasar KBBI yang digunakan pada algoritma *Levenshtein Distance* diperoleh dari *repository* GitHub yang bisa diakses pada tautan: <https://github.com/sastrawi/sastrawi/blob/master/data/kata-dasar.txt>

4. Model IndoBERT

BERT yang digunakan pada penelitian ini adalah model IndoBERT-base-phase-2 yang tersedia pada platform *HuggingFace* dan dapat diakses melalui tautan: <https://huggingface.co/indobenchmark/indobert-base-p2>

LEMBAR PERBAIKAN SKRIPSI

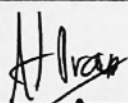


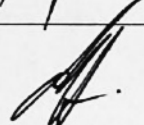
“NORMALISASI TEKS BERBASIS KAMUS SLANG DAN LEVENSHTAIN DISTANCE: IMPLEMENTASI PADA ANALISIS SENTIMEN EKSPEDISI SICEPAT EKSPRES DI TWITTER”

OLEH:

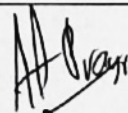
**DIKI SISWANTO
D42116316**

Skrripsi ini telah dipertahankan pada Ujian Akhir Sarjana tanggal 23 November 2022.
Telah dilakukan perbaikan penulisan dan isi skripsi berdasarkan usulan dari penguji dan pembimbing skripsi.

Persetujuan perbaikan oleh tim penguji:

	Nama	Tanda Tangan
Ketua	A. Ais Prayogi Alimuddin, S.T., M.Eng.	
Sekretaris	Anugrayani Bustamin, S.T., M.T.	
Anggota	Dr. Ir. Ingrid Nurtanio, M.T.	
	Ir. Christoforus Yohannes, M.T.	

Persetujuan perbaikan oleh pembimbing:

Pembimbing	Nama	Tanda Tangan
I	A. Ais Prayogi Alimuddin, S.T., M.Eng.	
II	Anugrayani Bustamin, S.T., M.T.	