

ENSEMBLE DEEP LEARNING UNTUK MENDETEKSI SARKASME PADA DATASET NEWS HEADLINE

Oleh:

Siti Nur Haliza

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

2022

Pendahuluan

- Bahasa terbentuk melalui serangkaian aturan, kaidah, atau pola yang ditetapkan untuk mengatur tata bunyi, bentuk kata, serta susunan kalimat. Pelanggaran terhadap aturan, kaidah, atau pola ini dapat mengganggu komunikasi[1]. Gaya bahasa memainkan peran penting dalam mengartikan maksud dan tujuan seseorang. Bahasa bisa berupa kata kasar yang penuh dengan sindiran dan mengandung olokan bisa menyakiti hati seseorang atau bisa diartikan sebagai sarkasme. Sarkasme dapat didefinisikan sebagai “ekspresi atau ucapan yang tajam, pahit, atau tajam; ejekan atau ejekan yang pahit” [2]. Menganalisis kalimat sarkasme bukan suatu hal yang mudah untuk dilakukan. Kalimat sarkasme tak jarang sekali disampaikan seseorang dalam menginformasikan sesuatu melalui sosial media.

Pendahuluan

- Deteksi sarkasme sering dijadikan untuk analisis sentimen dalam mengidentifikasi sentimen negatif, positif, atau netral. Sarkasme terdapat dalam ekspresi wajah, gerak tubuh, dan bahkan dalam teks [3]. Suatu kegagalan yang terjadi saat mendeteksi sarkasme dapat mempengaruhi hasil dalam klasifikasi analisis sentimen dari dataset. Tantangan sarkasme dan manfaat deteksi sarkasme untuk analisis sentimen telah menyebabkan minat penelitian deteksi sarkasme otomatis [4]. Berdasarkan hasil penelitian Alita dan Rahman [5], berhasil meningkatkan akurasi analisis sentimen sebesar 16,61% dengan mendeteksi sarkasme pada tweet tentang layanan publik. Dalam penelitian yang telah dilakukan oleh Saifullah dkk [6] yang melakukan deteksi sarkasme pada sosial media twitter. Sebanyak 17% pengguna sosial media Instagram menggunakan kalimat sarkasme, menunjukkan bahwa penggunaan kalimat sarkasme masih sering ditemukan pada platform tersebut.

Pendahuluan

- Pada penelitian yang telah disebutkan sebelumnya, digunakan metode *base classifier* machine learning dan deep learning dalam proses klasifikasi. Belum ada penelitian yang menggunakan ensemble deep learning untuk mendeteksi sarkasme. Berdasarkan hal tersebut, ada peluang untuk pengembangan dalam deteksi sarkasme menggunakan metode ensemble deep learning.
- Penelitian ini bertujuan untuk membangun model ensemble deep learning. Dalam penelitian ini, metode deep learning dasar yang akan digunakan meliputi Bidirectional Gated Recurrent Unit (BiGRU) dan Convolutional Neural Network (CNN). Pada penelitian ini, LightGBM digunakan untuk melakukan ensemble metode deep learning. Setelah melalui proses uji coba, diharapkan dapat ditemukan metode deep learning yang memiliki performa terbaik, sehingga dapat dijadikan acuan dalam pembuatan model untuk mendeteksi sarkasme pada judul berita di platform media sosial.

Rumusan dan Batasan Masalah

Rumusan Masalah

Rumusan masalah yang menjadi topik utama penelitian ini “Bagaimana komparasi antara metode Bidirectional Gated Recurrent Unit (BiGRU), Convolutional Neural Network (CNN), dan ensemble LightGBM?.

Batasan Masalah

1. Dataset diambil dari situs web : <https://www.kaggle.com/rmisra/news-headlines-dataset-for-sarcasm-detection>. Berbahasa inggris.
2. Metode yang digunakan untuk klasifikasi adalah Bidirectional Gated Recurrent Unit (BiGRU), Convolutional Neural Network (CNN), dan ensemble LightGBM.
3. Jumlah data training yang digunakan adalah 20.431 data dan data testing berjumlah 4.007
4. Dataset sudah melalui tahap preprocessing data

Tujuan dan Manfaat

Tujuan

1. Membandingkan metode Bidirectional Gated Recurrent Unit (BiGRU), Convolutional Neural Network (CNN), dan ensemble LightGBM
2. Mengukur akurasi metode Bidirectional Gated Recurrent Unit (BiGRU), Convolutional Neural Network (CNN), dan ensemble LightGBM

Manfaat

- a. Menambah referensi khususnya yang berhubungan dengan analisis sentimen terkait studi kasus deteksi sarkasme.
- b. Sebagai pedoman untuk penelitian selanjutnya.

Penelitian Sebelumnya

Alita & Isnain, 2020

Judul : Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier

Pada riset tersebut melakukan tahapan ekstraksi fitur dengan 4 set fitur yaitu sentiment related, punctuation-relate, lexical and syntactic, dan pattern-relate dan diklasifikasikan dengan menggunakan metode Random Forest Classifier. Hasil pengujian tersebut masih memiliki akurasi yang rendah.

Ranti & Girsang, 2020

Judul : Indonesian Sarcasm Detection Using Convolutional Neural Network

Pada riset tersebut mengidentifikasi sarkasme dalam bahasa indonesia. Penelitian ini menghasilkan skor F1 90% pada 1138 komentar pada data yang diuji.

Penelitian Sebelumnya

Fanissa et al., 2018

Judul : Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking

Pada riset tersebut dilakukan proses analisis sentimen dilakukan dari preprocessing, seleksi fitur dengan metode Query Expansion Ranking, dan klasifikasi dengan Naive Bayes. Hasil seleksi fitur 75% memiliki akurasi terbaik sebesar 86.6%.

Jia & Zan, 2022

Judul : Context-Based Sarcasm Detection Model in Chinese Social Media Using BERT and Bi-GRU Models

Pada riset tersebut menggabungkan dua metode untuk deteksi sarkasme, sehingga dapat meningkatkan akurasi deteksi sarkasme situasional.

Penelitian Sebelumnya

Fadli & Hidayatullah, 2019

Judul : Identifikasi Cyberbullying pada Media Sosial Twitter Menggunakan Metode LSTM dan BiLSTM

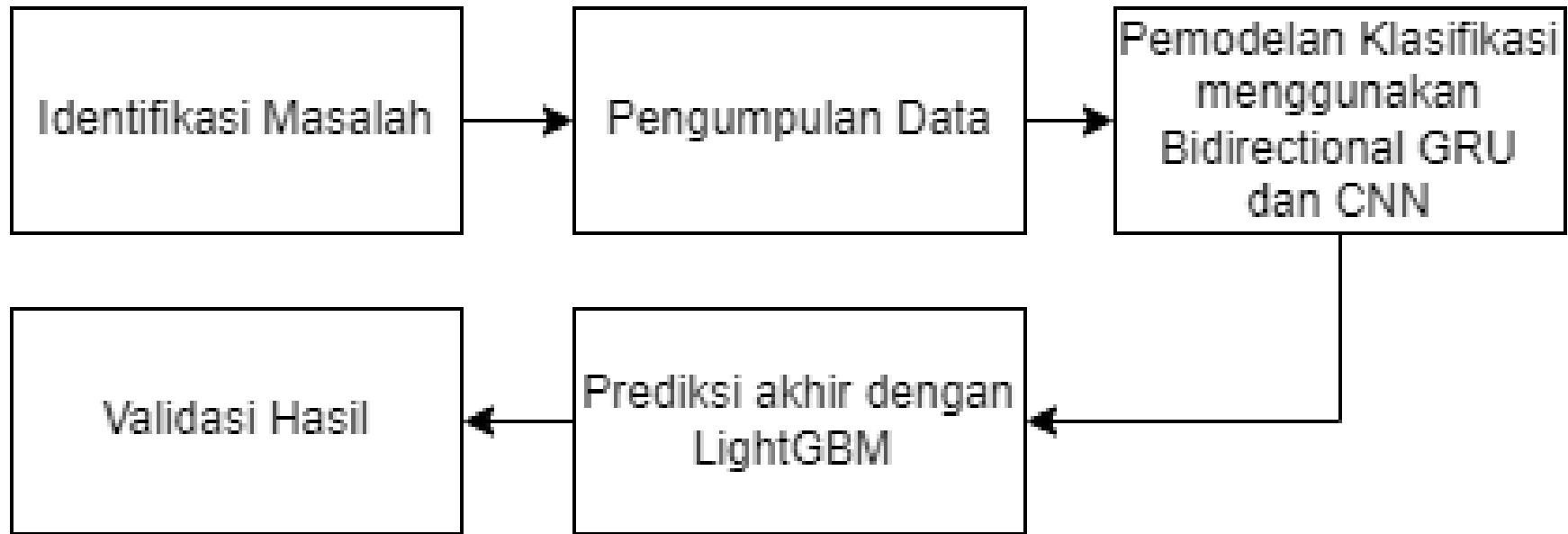
Pada riset tersebut menghasilkan akurasi dari metode LSTM 93.77% dan BiLSTM 95.24%. Penulis mengharapkan penelitiannya bisa dikembangkan dengan menambah dataset, menambah data training, menambah fitur kelas tidak hanya cyberbullying dan non.

Rosid et al., 2022

Judul : Pre-Trained Word Embeddings for Sarcasm Detection in Indonesian Tweets : A Comparative Study

Pada riset tersebut merepresentasikan vektor dari kata-kata, tiga model penyisipan kata pra-pelatihan yaitu GloVe, fastText, dan BERT. Hasil penelitian menunjukkan bahwa kombinasi fastText embeddings dan BiGRU sebagai pengklasifikasi menghasilkan performa terbaik dengan akurasi 93,85%.

Tahapan Penelitian



Pengumpulan Data

Pengumpulan data dilakukan dengan mencari di situs web yaitu Kaggle. Data yang diperoleh tersebut adalah data berita utama. Kumpulan data berita utama memiliki kesalahan tata bahasa yang minimal, kalimat yang unik dan data yang sangat bersih. Data tersebut menggunakan Natural Language Toolkit (NLTK) untuk tokenisasi, lemmatisasi dan pra pemrosesan. Dataset headline berita tersebut terdiri dari 26.709 headline berita.

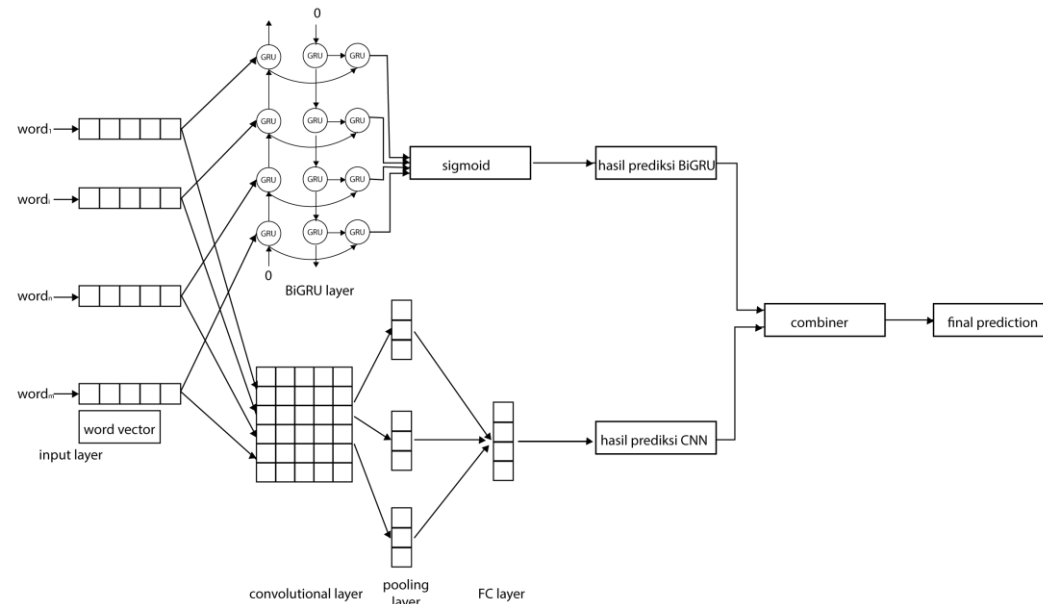
	headline	is_sarcastic
0	former versace store clerk sues over secret 'b...	0
1	the 'roseanne' revival catches up to our thorn...	0
2	mom starting to fear son's web series closest ...	1
3	boehner just wants wife to listen, not come up...	1
4	j.k. rowling wishes snape happy birthday in th...	0
5	advancing the world's women	0
6	the fascinating case for eating lab-grown meat	0
7	this ceo will send your kids to school, if you...	0
8	top snake handler leaves sinking huckabee camp...	1
9	friday's morning email: inside trump's presser...	0

Proses Klasifikasi Deep Learning

Proses klasifikasi diawali dengan mencari hasil dari Metode Bidirectional Gated Recurrent Unit dan Convolutional Neural Network. Hasil klasifikasi dari kedua metode kemudian digabungkan dengan algoritma LightGBM. Penelitian ini mengusulkan metode BiGRU sebagai alat ekstraksi fitur teks yang optimal. Parameter seperti Learning Rate, Batch Size, dan Epoch diukur dalam penelitian ini. Proses ekstraksi fitur dilakukan dengan menggunakan metode BiGRU yang melakukan ekstraksi kata secara forward dan backward melalui dua arah, yaitu layer forward yang mengambil kata dari awal hingga akhir, dan backward yang mengambil kata dari akhir hingga awal. Algoritma CNN memiliki keunggulan dalam hal kecepatan proses, dimana CNN terdapat layer konvolusi dan layer pooling yang dapat mereduksi dimensi data.

Prediksi Akhir

Prediksi akhir menggunakan algoritma LightGBM. LightGBM merupakan salah satu algoritma ansambel. Prediksi akhir ini dilakukan setelah hasil akurasi penilaian dan F1 score yang dihitung menggunakan instrument yang disebut confusion matrix dari klasifikasi dengan metode BiGRU dan CNN ditemukan, karena data prediksi akhir dengan algoritma LightGBM diambil dari hasil metode BiGRU dan CNN.



Prediksi Akhir

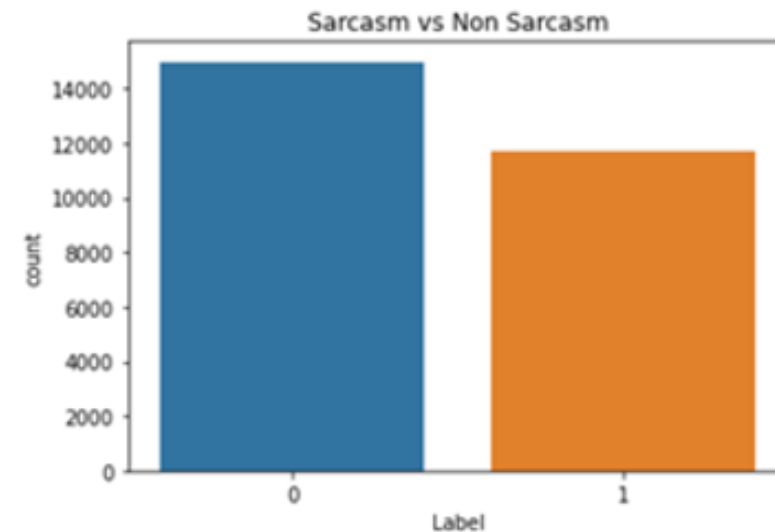
Prediksi akhir menggunakan algoritma LightGBM. LightGBM merupakan salah satu algoritma ansambel. LightGBM mengimplementasikan algoritma Gradient Boosting Decision Tree (GBDT) konvensional dengan penambahan dua teknik baru yakni Pengambilan Sampel Satu Sisi Berbasis Gradien (GOSS) dan Bundling Fitur Eksklusif (EFB). Teknik-teknik ini dirancang untuk secara signifikan meningkatkan efisiensi dan skalabilitas GBDT. Model LightGBM yang diusulkan menyediakan sistem pendukung keputusan yang dioptimalkan untuk pengguna. Kekhasan dari pendekatan yang diusulkan adalah dalam prosedur yang digunakan untuk menghitung jumlah pohon keputusan, kedalaman maksimum pohon, dan jumlah daun pohon untuk membangun model LightGBM yang optimal. Prediksi akhir ini dilakukan setelah hasil akurasi penilaian dan F1 score yang dihitung menggunakan instrument yang disebut confusion matrix dari klasifikasi dengan metode BiGRU dan CNN ditemukan, karena data prediksi akhir dengan algoritma LightGBM diambil dari hasil metode BiGRU dan CNN.

Hasil dan Pembahasan

Penelitian ini menggunakan dataset sebanyak 26.709 headline berita yang terdiri dari 11.724 yang bermakna sarkasme dan 14985 yang tidak bermakna sarkasme. Sebanyak 20.431 data untuk training dan 4.007 data untuk testing. Percobaan data dimulai dari 500, 1000 dan 2000. Jumlah epoch yang dicoba adalah 50 dan 100.

Tabel 1. Perbandingan Data Terhadap Label Sentimen

Label	Jumlah Data
0 (sarkasme)	11724
1 (tidak sarkasme)	14985



Gambar 5. Diagram Penyebaran Dataset

Hasil dan Pembahasan

Epoch 50 - BiGRU				
No.	Epoch	Dimensi	Akurasi	F1-Score
1	50	500	84,7%	83,6%
2	50	1000	85,1%	84,6%
3	50	2000	84,7%	84,1%
Rata-rata			84,8%	84,1%

Epoch 50 - CNN				
No.	Epoch	Dimensi	Akurasi	F1-Score
1	50	500	86,7%	85,2%
2	50	1000	86,7%	85,4%
3	50	2000	86,3%	84,7%
Rata-rata			86,6%	85,1%

Epoch 50 - LightGBM				
No.	Epoch	Dimensi	Akurasi	F1-Score
1	50	500	91,5%	90,4%
2	50	1000	91,9%	90,8%
3	50	2000	91,8%	90,8%
Rata-rata			91,7%	90,7%

Epoch 100 - BiGRU				
No.	Epoch	Dimensi	Akurasi	F1-Score
1	100	500	84,2%	83,7%
2	100	1000	85,3%	84,7%
3	100	2000	84,3%	83,9%
Rata-rata			84,6%	84,1%

Epoch 100 - CNN				
No.	Epoch	Dimensi	Akurasi	F1-Score
1	100	500	86,7%	84,9%
2	100	1000	86,5%	85,0%
3	100	2000	86,0%	84,5%
Rata-rata			86,4%	84,8%

Epoch 100 - LightGBM				
No.	Epoch	Dimensi	Akurasi	F1-Score
1	100	500	91,2%	90,2%
2	100	1000	91,2%	90,2%
3	100	2000	90,8%	89,7%
Rata-rata			91,1%	90,0%

Hasil dan Pembahasan

Nilai rata-rata dari percobaan dimensi BiGRU, CNN, dan LightGBM terbaik diperoleh pada *epoch* 50. Percobaan metode terhadap dataset dilakukan untuk melihat nilai akurasi dan f1-score dari metode ensemble deep learning yang diusulkan. Metode yang diuji adalah BiGRU, CNN, dan ensemble LightGBM.

Tabel 8. Perbandingan dari metode BiGRU, CNN, dan LightGBM

Metode	Akurasi	F1-Score
LightGBM	91,2%	90,2%
CNN	86,7%	85,4%
BiGRU	85,1%	84,6%

Kesimpulan

Metode terbaik untuk mendeteksi sarkasme adalah menggunakan metode LightGBM dengan dimensi 1000, karena mampu memberikan hasil klasifikasi yang paling akurat dibandingkan metode lain, dengan nilai akurasi 91,2% dan f1-score 90,2%. Keakuratan tersebut disebabkan oleh dimensi 1000 yang dapat merepresentasikan sentimen dari data secara tepat. Pemilihan jumlah epoch yang tepat juga berpengaruh pada hasil ensemble LightGBM, yang pada penelitian ini menggunakan epoch 50. Namun, terlalu banyak epoch dapat menghasilkan fitur-fitur yang kurang penting dan menjadi noise. Pada penelitian selanjutnya, akan dilakukan pengujian metode ensemble deep learning pada dataset dengan modalitas yang berbeda.

Referensi

- Alita, D., & Isnain, A. R. (2020). Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier. *Jurnal Komputasi*, 8(2), 50–58. <https://doi.org/10.23960/komputasi.v8i2.2615>
- Bigru, C. N. N. (2022). *applied sciences Short Text Aspect-Based Sentiment Analysis Based on*.
- Fadli, H. F., & Hidayatullah, A. F. (2019). Identifikasi Cyberbullying Pada Media Sosial Twitter Menggunakan Metode Klasifikasi Random Forest. *Automata*. <https://journal.uui.ac.id/AUTOMATA/article/download/17364/10897>
- Fanissa, S., Fauzi, M. A., & Adinugroho, S. (2018). Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking. *Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(8), 2766–2770. <http://j-ptiik.ub.ac.id>
- Fikri, M. I., Sabrila, T. S., & Azhar, Y. (2020). Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter. *Smatika Jurnal*, 10(02), 71–76. <https://doi.org/10.32664/smatika.v10i02.455>
- Govindan, V., & Balakrishnan, V. (2022). A machine learning approach in analysing the effect of hyperboles using negative sentiment tweets for sarcasm detection. *Journal of King Saud University - Computer and Information Sciences*, 34(8), 5110–5120. <https://doi.org/10.1016/j.jksuci.2022.01.008>
- Heru, A. (2018). Gaya Bahasa Sindiran Ironi, Sinisme Dan Sarkasme Dalam Berita Utama Harian Kompas. *Jurnal Pembahsi (Pembelajaran Bahasa Dan Sastra Indonesia)*, 8(2), 43. <https://doi.org/10.31851/pembahsi.v8i2.2083>
- Jia, C., & Zan, H. (2022). Context-Based Sarcasm Detection Model in Chinese Social Media Using BERT and Bi-GRU Models. *CEUR Workshop Proceedings*, 3150, 42–50.
- Mehndiratta, P., Sachdeva, S., & Soni, D. (2017). Detection of sarcasm in text data using deep convolutional neural networks. *Scalable Computing*, 18(3), 219–228. <https://doi.org/10.12694/scpe.v18i3.1302>
- Poria, S., Cambria, E., Hazarika, D., & Vij, P. (2016). A deeper look into sarcastic tweets using deep convolutional neural networks. *COLING 2016 - 26th International Conference on Computational Linguistics, Proceedings of COLING 2016: Technical Papers*, 1601–1612.

Referensi

Ranti, K. S., & Girsang, A. S. (2020). Indonesian Sarcasm Detection Using Convolutional Neural Network. *International Journal of Emerging Trends in Engineering Research*, 8(9), 4952–4955. <https://doi.org/10.30534/ijeter/2020/10892020>

Razali, M. S., Halin, A. A., Ye, L., Doraisamy, S., & Norowi, N. M. (2021). Sarcasm Detection Using Deep Learning with Contextual Features. *IEEE Access*, 9, 68609–68618. <https://doi.org/10.1109/ACCESS.2021.3076789>

Rosid, M. A., Siahaan, D., & Saikhu, A. (2022). *Pre-Trained Word Embeddings for Sarcasm Detection in Indonesian Tweets: A Comparative Study*. 281–286. <https://doi.org/10.1109/icitacee55701.2022.9924084>

Rufo, D. D., Debelee, T. G., Ibenthal, A., & Negera, W. G. (2021). Diagnosis of diabetes mellitus using gradient boosting machine (Lightgbm). *Diagnostics*, 11(9), 1–14. <https://doi.org/10.3390/diagnostics11091714>

Siti Farmida, Ediwarman, E., & Tisnasari, S. (2021). Analisis Satire Dan Sarkasme Dalam Debat Capres 2019 Dan Implementasinya Terhadap Pembelajaran Di Sma. *Bahtera Indonesia; Jurnal Penelitian Bahasa Dan Sastra Indonesia*, 6(2), 189–202. <https://doi.org/10.31943/bi.v6i2.131>

