



Analisis Sentimen Peringkat Pengguna Aplikasi Layanan Online Transportasi Melalui Komentar di Twitter

Natalia Silalahi

Universitas Negeri Medan, Indonesia, email: natalia.novena.silalahi@gmail.com

Info Artikel

Kata Kunci:

Analisis; Twitter; Sentimen; Pengguna; Aplikasi Layanan Online

Keywords:

Analysis; Twitter; Sentiment; User; Online Service Application



Lisensi: cc-by-sa

Copyright © 2023 nama_penulis

Abstrak

Berkembangnya dunia teknologi informasi, transportasi berkembang dengan adanya jasa transportasi online. Saat ini penggunaan Jasa transportasi online sudah seperti kebutuhan, melakukan analisis sentimen terhadap Jasa transportasi online untuk mengerahui mendapatkan informasi bagaimana pendapat masyarakat terhadap Jasa transportasi online tersebut. Data positive digunakan harus merupakan data positive. Media yang digunakan untuk mengambil data yang berasal dari platform media sosial, Twitter. Tugas akhir ini bertujuan untuk menganalisis tanggapan masyarakat dengan menganalisis data dalam bentuk tweet, kemudian menggunakan metode naive Bayes classifier untuk mengklasifikasinya ke dalam kategori positif dan negatif. Berdasarkan sistem yang telah dibangun didapatkan hasil 88,60% sentimen positif dan 11,40% sentimen negatif dengan akurasi 86,80%. Ternyata tingkat sentimen positif dari tweet publik lebih besar dari pada tingkat sentimen negatif.

Abstract

The development of the world of information technology, transportation is developing with the existence of online transportation services. Currently, the use of online transportation services has become a necessity, carrying out sentiment analysis of online transportation services to find out information about what the public thinks about these online transportation services. The positive data used must be positive data. The media used to retrieve data comes from the social media platform, Twitter. This final project aims to analyze public responses by analyzing data in the form of tweets, then using the naive Bayes classifier method to classify them into positive and negative categories. Based on the system that has been built, the results were 88.60% positive sentiment and 11.40% negative sentiment with an accuracy of 86.80%. It turns out that the level of positive sentiment from public tweets is greater than the level of negative sentiment.

1. PENDAHULUAN

Transportasi adalah penunjang penting bagi kehidupan sehari-hari. Seperti angkutan kota, taxi, dan pengojek masih digunakan sebagian masyarakat sebagai alat transportasi menuju ke tempat yang diinginkan [1]–[3]. Transportasi mulai berkembang dengan adanya Transportation online harganya yang relatif murah dibandingkan dengan conventional transportatio [4], [5]. Manfaat terbesar dari transportasi umum yang terus membaik akan sangat membantu mengurangi kemacetan jalan raya, polusi, serta mengurangi konsumsi bahan bakar dan energy [1], [6].

Dengan mendaftarkan diri sebagai sopir Transportation online baik kendaraan pribadi roda dua dan roda empat, kendaraan tersebut dapat bermanfaat sebagai conventional transportation yang bisa di order oleh masyarakat umum, memanfaatkan akses teknologi smartphone [7]–[9]. Selain itu aplikasi Transportation online tidak hanya dijadikan alat Transportation bagi masyarakat tapi juga bisa sebagai jasa gosend (pengirim barang) dan gofood (food delivery services) [1], [10]. Transportasi online terkenal karena memadukan teknologi komunikasi online untuk memberikan inovasi terkini dalam transportasi, sehingga masyarakat dapat dengan mudah memesan transportasi kapanpun dan dimanapun [11], [12].

Dengan fenomena tersebut, banyak masyarakat yang mengungkapkan pandangannya tentang transportasi online di Indonesia melalui media social [13]. Media sosial adalah layanan yang secara

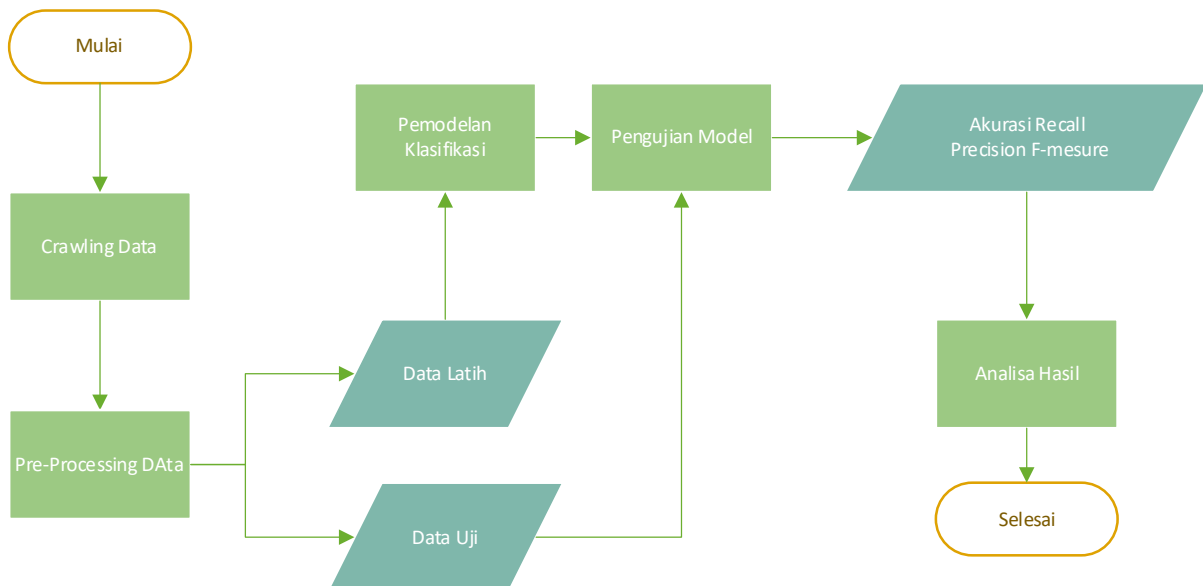
terus menerus mempromosikan pertukaran informasi dan topik dalam cakupan yang luas. Twitter adalah salah satu media sosial terpopuler di masyarakat. Twitter merupakan media sosial yang menghubungkan banyak orang dengan berbagai topik dari seluruh dunia. Dengan menggunakan Twitter, orang dapat mengungkapkan pendapatnya secara real time. Ini didukung oleh angka orang yang sudah memiliki ponsel dapat mengakses Internet dengan lebih mudah. Gunakan media sosial. Jumlah penduduk Indonesia saat ini adalah 262 juta. Pada 2017, lebih dari 50% orang (sekitar 143 juta) telah terhubung ke jaringan Internet. Dalam menganalisis sentimen publik di Twitter untuk memberikan informasi tentang kepuasan publik terhadap layanan lalu lintas online di Indonesia. Metode yang digunakan adalah naive Bayes classifier. Metode ini dipilih karena berbagai kelebihannya, termasuk kesederhanaan, kecepatan, dan akurasi yang tinggi. Data tersebut harus melalui tahap pra-pemrosesan sebelum diklasifikasikan. Setelah klasifikasi, Anda akan mendapatkan tweet positif atau negatif.

Dengan berkembangnya layanan transportasi online, orang sering membicarakannya. Masyarakat memberikan wawasannya melalui berbagai media, salah satunya media sosial Twitter. Masyarakat memiliki pendapat yang berbeda tentang layanan transportasi online. Di Twitter, perusahaan pelayanan online tersebut memiliki akun resmi yang dapat memberikan informasi mengenai layanan dan menampung komentar tweet dari masyarakat dan pelanggan. Menurut penelitian PeerReach, Indonesia merupakan pengguna Twitter paling aktif ketiga di dunia, yang berarti pengguna Twitter di Indonesia merupakan salah satu pengguna paling aktif di dunia. Jika kita memeriksa lebih lanjut tweet publik, kita dapat menganalisis sentimen.

Analisis sentimen adalah analisis peristiwa berdasarkan sikap masyarakat terhadap objek. Analisis sentimen biasanya dilakukan untuk mengumpulkan dan menemukan opini publik di postingan blog, Twitter, Facebook, dll. Analisis sentimen diperlukan untuk mengetahui opini publik terhadap suatu objek tertentu. Opini tersebut bisa negatif atau positif, tergantung opini publik terhadap objeknya. Oleh karena itu dalam penelitian ini perlu dilakukan analisis terhadap pendapat tersebut agar dapat dijadikan patokan atas jasa pengiriman online baik buruknya berdasarkan pelanggan

2. METODE PENELITIAN

Berikut beberapa tahapan metode penelitian:



Gambar 1. Metode Penelitian

2.1 Analisis Sentimen

Analisis sentimen adalah pemrosesan bahasa alami yang digunakan untuk menentukan emosi orang terhadap produk atau topik tertentu. Ada banyak metode untuk menganalisis sentimen media sosial Twitter, dan banyak penelitian telah mengembangkan metode yang ada menjadi metode presisi tinggi baru [9], [14]. Bandingkan beberapa metode, seperti metode Naive Bayes, metode K-nearest neighbour, serta kombinasi K-means dan LVQ. Dalam proses pengklasifikasian buku berbahasa Indonesia menggunakan data 200 buku digunakan 150 buku sebagai data latih. 50 buku digunakan sebagai data pelatihan. Digunakan sebagai data uji. Berdasarkan hasil penelitian akurasi dari metode

KNN adalah 96%, kemudian naive bayes 98%, kemudian kombinasi K-Means dan LVQ menghasilkan akurasi 92,2%. Metode Naive Bayes memperoleh hasil akurasi tertinggi [14].

2.2 Text Mining

Di Twitter kita bisa menuliskan status atau pesan untuk dibaca pengguna lain nanti, fitur ini disebut tweet [3], [5], [15]. Keunikan Twitter adalah mempublikasikan tweet atau postingan di Twitter, dengan ukuran maksimal 140 karakter [15] Salah satu kelebihan media sosial ini adalah Twitter menyediakan API (application programming interface) yang sangat baik sehingga memudahkan setiap orang untuk mengambil data dari Twitter [16]. Mengumpulkan data dari Twitter dapat digunakan untuk berbagai kebutuhan, seperti memahami popularitas kandidat kampanye, memperoleh informasi tentang popularitas produk atau hanya melihat semua mention, memposting ulang ke akun Twitter tertentu. Oleh karena itu, perusahaan semakin tertarik untuk mendapatkan informasi tentang pandangan dan perasaan masyarakat tentang produk dan jasanya melalui media sosial, salah satunya Twitter [17], [18]. Dalam penelitian ini, data dari Twitter merupakan data tentang opini pelanggan pengiriman online Gojek dan Grab.

2.3 Crawling Data

Crawling data di Twitter merupakan proses mengambil atau mengunduh data dari server Twitter berupa data pengguna dan data tweet dengan bantuan application programming interface (API) Twitter [16], [19]. Penjelajahan data ini dilakukan untuk mengambil data dari Twitter yang diperlukan untuk tugas akhir ini [4]. Cara pengambilan data adalah dengan membuat program dengan memasukkan kata kunci untuk mencari tweet yang diinginkan. Misalnya program "@gojekindonesia" menambahkan tweet yang menyebutkan pesan tersebut ke akun Gojek. Kumpulan tweet adalah data yang akan digunakan.

2.4 Pre-processing Teks

Karena masih banyak simbol dan kata yang tidak diperlukan, data yang diperoleh dari hasil grabbing tidak dapat diklasifikasi secara langsung, sehingga perlu dilakukan preprocess data untuk mengolah data yang lebih terstruktur dan bersih sehingga bisa diklasifikasikan [9], [20]. Text preprocessing adalah memproses data yang diperoleh melalui tahapan case folding, tokenisasi, filtering dan stemming [14].

2.5 Naïve Bayes Classifier

Klasifikasi Naive Bayes adalah metode klasifikasi statistik yang dapat memprediksi probabilitas anggota kelas, seperti probabilitas bahwa sampel yang diberikan termasuk dalam kelas tertentu [21]. Penelitian ini akan menggunakan metode ini untuk mengklasifikasikan data yang diekstrak dari Twitter dan membagi data tersebut menjadi kategori positif dan [2], [22], [23]. Klasifikasi Naive Bayes adalah metode klasifikasi berdasarkan teorema Bayes, yang memprediksi peluang masa depan berdasarkan pengalaman masa lalu [24]. Fitur utama dari pengklasifikasi Naive Bayes adalah hipotesis yang sangat kuat untuk setiap situasi atau peristiwa. Metode ini sangat cocok untuk klasifikasi sentimen dalam tugas akhir ini karena memiliki beberapa keunggulan diantaranya kesederhanaan, kecepatan dan akurasi yang tinggi.

2.6 Twitter

Di Twitter kita bisa menuliskan status atau pesan untuk dibaca pengguna lain nanti, fitur ini disebut tweet [15]. Keunikan Twitter adalah mempublikasikan tweet atau postingan di Twitter, dengan ukuran maksimal 140 karakter [25]. Salah satu kelebihan media sosial ini adalah Twitter menyediakan API (application programming interface) yang sangat baik sehingga memudahkan setiap orang untuk mengambil data dari Twitter [16]. Mengumpulkan data dari Twitter dapat digunakan untuk berbagai kebutuhan, seperti memahami popularitas kandidat kampanye, memperoleh informasi tentang popularitas produk atau hanya melihat semua mention, memposting ulang ke akun Twitter tertentu. Oleh karena itu, perusahaan semakin tertarik untuk mendapatkan informasi tentang pandangan dan perasaan masyarakat tentang produk dan jasanya melalui media sosial, salah satunya Twitter [16], [17]. Dalam penelitian ini, data dari Twitter merupakan data tentang opini pelanggan pengiriman online Gojek dan Grab.

3. HASIL DAN ANALISIS

Pengujian mode dihasilkan setelah menandai data kombinasi kedua akun tersebut sebagai data, maka data tersebut dibagi menjadi data latih dan data uji acak. Pembagian data pelatihan meliputi 750

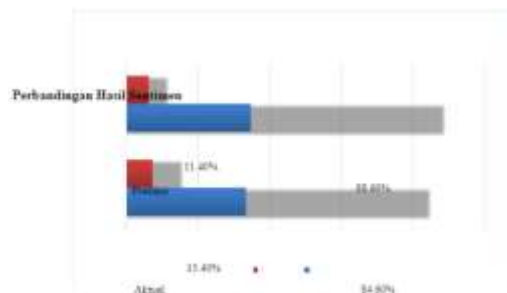
data akun @gojekindonesia dan 750 data akun @grabID. Data pengujian berisi 250 record untuk setiap akun. Data pelatihan digunakan untuk melaksanakan model pelatihan yang telah dirancang sehingga program dapat memahami apa yang ingin dicapai dari penelitian ini. Data pelatihan mencakup tweet dan tag yang diselesaikan secara manual. Data pengujian digunakan sebagai determinan model pengklasifikasi pengujian, dan setiap data diprediksi berdasarkan dua kategori positif dan negatif. Data uji yang digunakan sebanyak 500 data. Kemudian model yang ada digunakan untuk memprediksi setiap kata yang telah dipisahkan. Proses prediksi kelas dilakukan menggunakan pustaka Scikit-learn dalam bahasa pemrograman python acberdasarkan metode naive Baye s.

Table 1

Tabel 1. Hasil Pengujian Model

Kelas	Prediksi	Prediksi	Total	Recall	Precision	F-measure
	Positif	Negatif		Recall	Precision	F-measure
Aktual	400	23	423	90.29%	94.56%	92.38%
Positif	400	23	423			
Aktual	43	34	77	90.29%	94.56%	92.38%
Negatif	43	34	77			
Akurasi			500	86.80%		

Berdasarkan Tabel 1 (hasil tes model dijelaskan dengan membandingkan kelas prediksi dan kelas sebenarnya), tingkat recall dihitung sebesar 90,29%, dan keakuratan dihitung sebagai 94,56%. Dan perhitungan f-measure 9 tingkat. Kemudian, hasil prediksi diperoleh dari data uji yang digunakan untuk menguji model. Akurasi model prediksi yang digunakan adalah 86.80%. Berdasarkan cuitan tentang transportasi online di Indonesia, diperoleh pula hasil prediksi sentimen positif sebesar 88,60% dan hasil prediksi sentimen negatif sebesar 11,40%. Hasil prediksi sentimen dapat dilihat pada grafik pada Gambar 4.



Gambar 1. Perbandingan hasil sentimen

Analisis hasil pengujian terkait dengan proses evaluasi menggunakan 1500 data latih dan 500 data uji. Selama pengujian data pengujian, hasilnya ditunjukkan pada Tabel 2.

Tabel 2. Contoh perbedaan prediksi model

No	Tweet	Kelas	Prediksi
		Aktual	Model
1	dear gojekindonesia maps lokasi driver berbeda kenyataannya	0	1
2	govejindonesia suka tampilan aplikasinya terima kasih	1	1
3	cc grabid tolong ditegur driver anarkis	0	0

Dapat dilihat dari Tabel 2 bahwa hasil evaluasi prediksi sentimen masih dapat memprediksi kesalahan dalam sistem, sehingga akurasi program lebih rendah dari keadaan optimal.

4. KESIMPULAN

Berdasarkan hasil penggunaan metode Naive Bayes untuk menganalisis sentimen jasa transportasi online untuk menentukan kategori prediksi, sehingga diperoleh hasil emosional yang positif. 88,60%, sentimen negatif 11,40%, akurasi 86,80%. Hasil penelitian menunjukkan bahwa tingkat emosi positif lebih besar daripada tingkat emosi negatif. Berdasarkan hasil penelitian dapat diketahui bahwa

menurut data, banyak masyarakat di Indonesia yang puas dengan layanan transportasi online. Namun penelitian ini masih memiliki beberapa kekurangan. Hal ini mungkin terjadi karena kurangnya fitur yang digunakan dalam data pelatihan, dan hasil pelabelan emosional dari data tersebut berbeda (perbedaan dalam data pengujian manual). Hasil prediksi sentiment diperoleh dari hasil klasifikasi model. Proses membangun fitur dan pelabelan data merupakan dasar untuk membuat model yang baik. Label masih subyektif dan kurang objektif

REFERENSI

- [1] E. Putri and F. M. Sari, "INDONESIAN EFL STUDENTS' PERSPECTIVES TOWARDS LEARNING MANAGEMENT SYSTEM SOFTWARE," *Journal of English Language Teaching and Learning*, vol. 1, no. 1, pp. 20–24, Jun. 2020, doi: 10.33365/jeltl.v1i1.244.
- [2] A. P. Zanofa, R. Arrahman, M. Bakri, and A. Budiman, "PINTU GERBANG OTOMATIS BERBASIS MIKROKONTROLER ARDUINO UNO R3," *Jurnal Teknik dan Sistem Komputer*, vol. 1, no. 1, pp. 22–27, Jun. 2020, doi: 10.33365/jtikom.v1i1.76.
- [3] S. Samanik and F. Lianasari, "Antimatter Technology: The Bridge between Science and Religion toward Universe Creation Theory Illustrated in Dan Brown's Angels and Demons," *TEKNOSASTIK*, vol. 14, no. 2, p. 18, Apr. 2018, doi: 10.33365/ts.v14i2.58.
- [4] R. Indriani and M. A. Firmansyah, "STRATEGI KOMUNIKASI PEMASARAN MELALUI BTS MEAL OLEH RESTORAN MC.DONALDS DAN PERSEPSI KONSUMEN (Studi Gerai Mc. Donald's Kota Bengkulu)," *JOISCOM (Journal of Islamic Communication)*, vol. 3, no. 1, pp. 1–12, Apr. 2022, doi: 10.36085/joiscom.v3i1.3315.
- [5] D. Hariyanto and F. Prasetyo, "Penerapan Metode Unified Approach Pada Sistem Informasi Akademik Berbasis Website," *Jurnal Khatulistiwa Informatika*, vol. 7, no. 1, Jul. 2019, doi: 10.31294/jki.v7i1.5739.
- [6] N. P. Wulandari, "USE OF ENGLISH LEARNING ACCOUNTS IN INSTAGRAM TO IMPROVE EFL STUDENTS VOCABULARY," *ENGLISH JOURNAL OF INDRAGIRI*, vol. 6, no. 2, pp. 363–374, Jul. 2022, doi: 10.32520/eji.v6i2.2009.
- [7] T. I. Setri and D. B. Setiawan, "Matriarchal Society in The Secret Life of Bees by Sue Monk Kidd," *Linguistics and Literature Journal*, vol. 1, no. 1, pp. 28–33, Jun. 2020, doi: 10.33365/lj.v1i1.223.
- [8] Y. Mertania and D. Amelia, "Black Skin White Mask: Hybrid Identity of the Main Character as Depicted in Tagores The Home and The World," *Linguistics and Literature Journal*, vol. 1, no. 1, pp. 7–12, Jun. 2020, doi: 10.33365/lj.v1i1.233.
- [9] M. Lestari and A. Y. Wahyudin, "LANGUAGE LEARNING STRATEGIES OF UNDERGRADUATE EFL STUDENTS," *Journal of English Language Teaching and Learning*, vol. 1, no. 1, pp. 25–30, Jun. 2020, doi: 10.33365/jeltl.v1i1.242.
- [10] D. Apriyanti and M. Ayu, "THINK-PAIR-SHARE: ENGAGING STUDENTS IN SPEAKING ACTIVITIES IN CLASSROOM," *Journal of English Language Teaching and Learning*, vol. 1, no. 1, pp. 13–19, Jun. 2020, doi: 10.33365/jeltl.v1i1.246.
- [11] S. Suprayogi, S. Samanik, and E. P. Chaniago, "Penerapan Teknik Mind Mapping, Impersonating dan Questioning dalam Pembelajaran Pidato di SMAN 1 Semaka," *JAMU: Jurnal Abdi Masyarakat UMUS*, vol. 2, no. 01, pp. 33–40, Aug. 2021, doi: 10.46772/jamu.v1i02.475.
- [12] M. A. Firmansyah, S. Karlinah, and S. Sumartias, "Kampanye Pilpres 2014 dalam Konstruksi Akun Twitter Pendukung Capres," *Jurnal The Messenger*, vol. 9, no. 1, p. 79, Jan. 2017, doi: 10.26623/themessenger.v9i1.430.
- [13] M. Fithratullah, "Representation of Korean Values Sustainability in American Remake Movies," *TEKNOSASTIK*, vol. 19, no. 1, p. 60, Jan. 2021, doi: 10.33365/ts.v19i1.874.
- [14] A. R. Utami, B. Yulianto, and R. K. Agustina, "Fostering Students' Speaking Ability through Traditional Talking Stick," *English Language in Focus (ELIF)*, vol. 2, no. 2, p. 71, Apr. 2020, doi: 10.24853/elif.2.2.71-78.
- [15] T. Wulandari and S. Samanik, "Dissociative Identity Disorder And Its Significance To Nina Sayers Personality Development," *Journal of Arts and Education*, vol. 3, no. 1, Jan. 2023, doi: 10.33365/jae.v3i1.200.
- [16] R. N. N. Az-zahra, M. P. Tyo, F. Rahman, M. Z. Sidiq, and R. Sanjaya, "Analisis Kualitas Layanan Terhadap Kepuasan Pengguna Sistem Informasi Kawasan Agrowisata Cicantayan Menggunakan Webqual 4.0," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 8, no. 1, p. 13, Apr. 2022, doi: 10.26418/jp.v8i1.51056.
- [17] H. Sulistiani and A. A. Aldino, "DECISION TREE C4.5 ALGORITHM FOR TUITION AID GRANT PROGRAM CLASSIFICATION (CASE STUDY: DEPARTMENT OF INFORMATION SYSTEM, UNIVERSITAS TEKNOKRAT INDONESIA)," *Eduatic - Scientific Journal of Informatics Education*, vol. 7, no. 1, Nov. 2020, doi: 10.21107/edutic.v7i1.8849.
- [18] M. Fithratullah, "Globalization and Culture Hybridity The Commodification on Korean Music and its Successful World Expansion," *Digital Press Social Sciences and Humanities*, vol. 2, p. 13, 2019, doi: 10.29037/digitalpress.42264.
- [19] Y. Hartanto, M. A. Firmansyah, and L. Adhrianti, "Implementation Digital Marketing Pesona 88 Curup in to Build Image for the Decision of Visit Tourist Attraction," in *Advances in Social Science, Education and Humanities Research*, Atlantis Press, 2022. doi: 10.2991/assehr.k.220407.121.

- [20] N. H. Siregar and M. Siregar, "THE ANALYSIS OF THE READING MATERIALS IN ENGLISH ON SKY 2 TEXTBOOK BASED ON SCHOOL-BASED CURRICULUM FOR GRADE VIII STUDENTS OF JUNIOR HIGH SCHOOL," *GENRE Journal of Applied Linguistics of FBS Unimed*, vol. 6, no. 4, Jan. 2018, doi: 10.24114/genre.v6i4.8531.
- [21] Y. Gustanti and M. Ayu, "THE CORRELATION BETWEEN COGNITIVE READING STRATEGIES AND STUDENTS' ENGLISH PROFICIENCY TEST SCORE," *Journal of English Language Teaching and Learning*, vol. 2, no. 2, pp. 95–100, Dec. 2021, doi: 10.33365/jeltl.v2i2.1452.
- [22] V. A. Safitri, L. Sari, and R. R. Gamayuni, "Research and Development, Environmental Investments, to Eco-Efficiency, and Firm Value," *The Indonesian Journal of Accounting Research*, vol. 22, no. 03, Sep. 2019, doi: 10.33312/ijar.446.
- [23] S. D. Ramdan, "PENGEMBANGAN KOPER PINTAR BERBASIS ARDUINO," *Journal ICTEE*, vol. 1, no. 1, Sep. 2020, doi: 10.33365/jictee.v1i1.699.
- [24] M. P. K. Putra, "Deteksi Bola Multipola Memanfaatkan Ekstraksi Fitur Local Binary Pattern dengan Algoritma Learning Adaboost," *Journal of Engineering, Computer Science and Information Technology (JECSIT)*, vol. 2, no. 1, Mar. 2023, doi: 10.33365/jecsit.v1i1.1.
- [25] K. A. Dharlie and S. Samanik, "IMAGERY ANALYSIS IN MATSUOKA'S CLOUD OF SPARROWS," *Linguistics and Literature Journal*, vol. 2, no. 1, pp. 17–24, Jun. 2021, doi: 10.33365/lj.v2i1.514.