

MENINGKATKAN PERFORMA ULASAN BERBAHASA INDONESIA DENGAN SPELLING CORRECTOR PETER NORVIG DAN PELABELAN SENTISTRENGTH_ID

Yessy Asri[✉], Dwina Kuswardani, Josephine Ferdinanda Purba TS

Fakultas Telematika Energi, Institut Teknologi PLN, Jakarta, Indonesia

Email: yessyasri@itpln.ac.id

DOI: <https://doi.org/10.46880/jmika.Vol9No1.pp92-98>

ABSTRACT

The development of information technology drives digital transformation in various sectors, including public services. Evaluating the quality of mobile-based applications often relies on user reviews and ratings. However, the mismatch between the two can affect the understanding of user satisfaction. This research analyzes the sentiment of PLN Mobile application reviews using an approach that combines the lexicon-based SentiStrength_id method and the Support Vector Machine (SVM) model. Data was collected through web scraping from the Google Play Store and then processed through preprocessing stages, including spelling correction using the Peter Norvig method. The next process involved tokenization, stopword removal, and sentiment labeling using SentiStrength_id. The analysis results showed that the SVM model with a training and test data ratio of 90:10 achieved the highest accuracy of 82%. From 11,004 review data, the sentiment labeling results showed 67.4% positive, 16.5% negative, and 16.1% neutral. Compared to the users' initial ratings, a mismatch of 5% was found for the positive class (553 data), 6.6% for the negative (720 data), and 11.6% for the neutral (1,273 data). Word cloud analysis showed that positive sentiments were dominated by words such as "helpful," "easy," and "fast," while negative sentiments were related to "no," "problem," and "interference," reflecting the app's technical issues. Neutral sentiments contain words related to questions about services and electricity tokens. This research shows that the combination of Peter Norvig's method and SentiStrength_id successfully improves the quality of sentiment analysis. These results confirm the importance of optimal preprocessing in improving the accuracy of machine learning-based sentiment modeling.

Keyword: *Sentiment Analysis, PLN Mobile, SentiStrength_id, Support Vector Machine, Web Scraping, Peter Norvig.*

ABSTRAK

Perkembangan teknologi informasi mendorong transformasi digital di berbagai sektor, termasuk layanan publik. Evaluasi kualitas aplikasi berbasis mobile sering kali bergantung pada ulasan dan rating pengguna. Namun, ketidaksesuaian antara keduanya dapat mempengaruhi pemahaman terhadap kepuasan pelanggan. Penelitian ini menganalisis sentimen ulasan aplikasi PLN Mobile dengan menggunakan pendekatan yang menggabungkan metode SentiStrength_id berbasis leksikon dan model Support Vector Machine (SVM). Data dikumpulkan melalui teknik web scraping dari Google Play Store dan kemudian diproses melalui tahapan preprocessing, termasuk koreksi ejaan dengan metode Peter Norvig. Proses selanjutnya adalah tokenisasi, penghilangan stopword, dan pelabelan sentimen menggunakan SentiStrength_id. Hasil analisis menunjukkan bahwa model SVM dengan rasio data latih dan data uji sebesar 90:10 mencapai akurasi tertinggi sebesar 82%. Dari 11.004 data ulasan, hasil pelabelan sentimen menunjukkan 67,4% positif, 16,5% negatif, dan 16,1% netral. Dibandingkan dengan penilaian awal pengguna, ditemukan ketidaksesuaian sebesar 5% untuk kelas positif (553 data), 6,6% untuk kelas negatif (720 data), dan 11,6% untuk kelas netral (1.273 data). Analisis word cloud menunjukkan bahwa sentimen positif didominasi oleh kata-kata seperti "membantu", "mudah", dan "cepat", sedangkan sentimen negatif terkait dengan "tidak", "masalah", dan "gangguan", yang mencerminkan masalah teknis aplikasi. Sentimen netral berisi kata-kata yang berkaitan dengan pertanyaan tentang layanan dan token listrik. Penelitian ini menunjukkan bahwa kombinasi metode Peter Norvig dan SentiStrength_id berhasil meningkatkan kualitas analisis sentimen. Hasil ini menegaskan pentingnya preprocessing yang optimal dalam meningkatkan akurasi pemodelan sentimen berbasis machine learning.

Kata Kunci: *Analisis Sentimen, PLN Mobile, SentiStrength_id, Support Vector Machine, Web Scraping, Peter Norvig.*

PENDAHULUAN

Perkembangan teknologi informasi telah mendorong peningkatan penggunaan internet dan ponsel seluler secara global, termasuk di Indonesia. Berdasarkan data dari Badan Pusat Statistik (BPS) tahun 2022, sebanyak 67,88% penduduk Indonesia menggunakan ponsel, sementara tingkat penggunaan internet mencapai 98,3% pada Januari 2023 (Ahdiat, 2023)(Annur, 2023). Peningkatan ini mendorong berbagai sektor terutama bidang pelayanan publik, untuk beralih dari sistem konvensional menuju digital. Salah satu bentuk transformasi tersebut dapat dilihat dari peningkatan pengguna aplikasi mobile yang dirancang untuk memperbaiki efisiensi dan produktivitas (Saefillah, 2023). Aplikasi mobile dapat diakses menggunakan ponsel atau tablet dan dapat diunduh dari platform seperti Google Play Store. Evaluasi kualitas dari aplikasi tersebut bergantung pada ulasan dan penilaian pengguna yang memberikan masukan untuk perbaikan. Ulasan biasanya berupa komentar terkait pengalaman pengguna, sementara penilaian dalam bentuk rating menunjukkan persepsi umum terhadap aplikasi (Sasmoko, 2020).

Namun, terdapat ketidaksesuaian yang kerap muncul antara ulasan dengan rating yang diberikan oleh pengguna. Beberapa ulasan dapat berisi kritik teknis namun tetap diiringi dengan rating tinggi, atau sebaliknya. Ketidaksesuaian ini dapat mempengaruhi evaluasi objektif kualitas aplikasi dan pada akhirnya berdampak pada pengembangan dan perbaikan aplikasi. Sebagai contoh, aplikasi PLN Mobile, yang dikembangkan oleh PT PLN (Persero) yang merupakan perusahaan kedua terbesar di Indonesia ("Fortune Indonesia 100," 2023), telah diunduh lebih dari 10.000.000 kali dan menerima lebih dari 807.000 ulasan dengan rating 4,8 (per Maret 2024). Aplikasi ini memiliki basis pengguna yang besar dan aktif, sehingga sangat relevan untuk dijadikan objek untuk menganalisis kepuasan dan kebutuhan pengguna. Oleh karena itu, mendapatkan gambaran yang jelas tentang ketidaksesuaian tersebut, analisis sentimen yang akurat sangat penting (Liu, 2010). Analisis sentimen menjadi krusial dalam mengevaluasi opini pengguna, yang dapat dilakukan menggunakan tiga pendekatan utama yaitu Machine Learning Approach, Knowledge Approach, dan Hybrid Approach yang menggabungkan kedua metode tersebut (M Isa, 2017). Pada penelitian ini diterapkan pendekatan hybrid dengan menggabungkan metode lexicon based menggunakan

SentiStrength_id dan model Support Vector Machine (SVM). Pemilihan model SVM didasarkan pada keunggulan akurasi dibandingkan metode lain seperti Naïve Bayes. SVM memiliki akurasi sebesar 81,46% lebih tinggi dibandingkan Naïve Bayes yang hanya mencapai 75,41% (Ilmawan & Mude, 2020). Pada proses analisis sentimen diperlukan proses pembobotan atau proses perubahan menjadi vector seperti TF-IDF yang dapat meningkatkan akurasi (Thomas & Rumaisa, 2022)(Handayani et al., 2020). Selain itu jumlah dataset juga mempengaruhi akurasi (Fatimah Rahmadayana & Yuliant Sibaroni, 2021).

Penelitian terkait pelabelan berbahasa Indonesia yang menggunakan SentiStrength_id. Penelitian pertama implementasi labelling SentiStrength terhadap teks berbahasa Indonesia yang menghasilkan akurasi sebesar 57,53% (Sianipar & Setiawan, 2015). Penelitian yang pertama kali mengadaptasi SentiStrength_id (Haryalesmana Wahid, 2016). Penerapan SentiStrength_id (Aulia et al., 2021)(Sari et al., 2021)(Khaira et al., 2020). Penelitian yang melakukan perbandingan, menyatakan bahwa SentiStrength_id dipilih karena akurasi lebih tinggi dibandingkan InSet. SentiStrength_id memiliki akurasi sebesar 64,46%, sementara InSet hanya mencapai 62,25% (Abdillah et al., 2021). Penelitian lain menyatakan SentiStrength_id unggul 82% (Ependi et al., 2023).

Selain hal tersebut, tantangan yang sering dihadapi dalam analisis sentimen adalah kesalahan pengetikan dalam ulasan yang dapat menurunkan ketepatan pelabelan sentimen. Oleh karena itu, penelitian ini menggunakan teknik spelling corrector, terdapat beberapa jenis dari spelling corrector (Simanjuntak et al., 2018)(Mutammiah et al., 2017). Selain itu terdapat perbandingan beberapa spelling corrector (Kusuma & Ratnasari, 2023)(Alfred & Teoh, 2019). Dari antara spelling corrector tersebut digunakan Peter Norvig pada tahap preprocessing untuk memperbaiki kesalahan ejaan yang sering muncul. Teknik ini sebelumnya terbukti efektif meningkatkan akurasi analisis sentimen hingga 90,03% (Malik Iryana & Pandu Adikara, 2021). Penelitian terkait PLN Mobile juga telah dilakukan menggunakan metode *lexicon-based Vader* dan Naïve Bayes dalam menganalisis ulasan aplikasi PLN Mobile dengan akurasi 70% (Asri et al., 2022). Namun penelitian tersebut menggunakan pelabelan dalam bahasa Inggris. Berbeda dengan penelitian tersebut, penelitian ini

menggunakan pelabelan berbahasan Indonesia menggunakan SentiStrength_id, serta menambahkan spelling corrector untuk meningkatkan ketepatan pelabelan dan menggunakan model Support Vector Machine untuk meningkatkan akurasi analisis sentimen ulasan pengguna aplikasi PLN Mobile.

METODE PENELITIAN

Metode penelitian ini terdiri dari beberapa tahapan, yaitu.

Alur Penelitian

Pada penelitian ini ditambahkan proses *spelling corrector* pada *preprocessing* menggunakan Peter Norvig, lalu dilakukan proses *labelling* menggunakan SentiStrength_id dan hasil pelabelan menggunakan SVM. Adapun alur penelitian ditunjukkan pada gambar 1 dibawah ini.



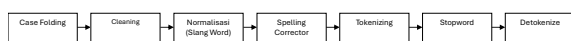
Gambar 1. Alur Penelitian

Dataset

Perolehan data yang akan digunakan pada penelitian ini menggunakan teknik web scrapping dari aplikasi PLN Mobile yang terdapat di platform Google Plays Store. Dari keseluruhan data yaitu 841.000 (Per 23 April 2024), peneliti melakukan scrapping dengan sample sebanyak 50.000 data ulasan terbaru. Lalu menyaring data untuk rentang waktu Januari 2022 hingga Desember 2023 dan diperoleh sebanyak 11.004 dataset.

Preprocessing

Setelah melakukan proses *web scrapping*, selanjutnya dilakukan proses pembersihan data yang nantinya akan dilabelkan (Kulkarni & Shivananda, 2019). Adapun proses *preprocessing* sebagai berikut:

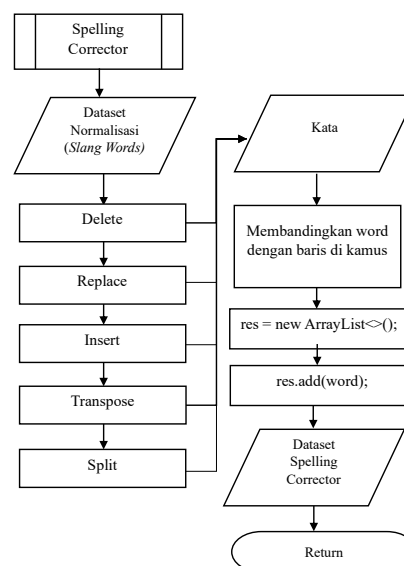


Gambar 2. Preprocessing

Proses case folding, cleaning (Handayani et al., 2020). Spelling Corrector digunakan untuk proses memperbaiki kata dengan mendeteksi kesalahan ketik dan memberikan rekomendasi kata yang benar (Martin et al., 2021). Tokenizing (Handayani et al., 2020), Stopword (Fatimah Rahmadayana & Yuliant Sibaroni, 2021). Pada proses stopwords ini akan diterapkan fungsi untuk tidak menghapus beberapa kata sesuai dengan kamus yang terdapat pada SentiStrength_id agar menghasilkan pelabelan yang maksimal. Detokenize

merupakan proses untuk mengembalikan ulasan yang sudah dibagi per kata menjadi kalimat, dikarenakan pelabelan sentistrength_id memiliki fungsi yang berkerja dalam bentuk string.

Pada penelitian ini, dilakukan penambahan Spelling Corrector menggunakan Peter Norvig dengan tujuan untuk memperbaiki kata yang mengalami kesalahan ejaan agar mendapatkan hasil sentimen yang tepat. Berikut adalah alur dari Peter Norvig:



Gambar 3. Proses Spelling Corrector Peter Norvig

$$PN = \operatorname{argmax} \frac{N(c)}{N} \quad (1)$$

Keterangan:

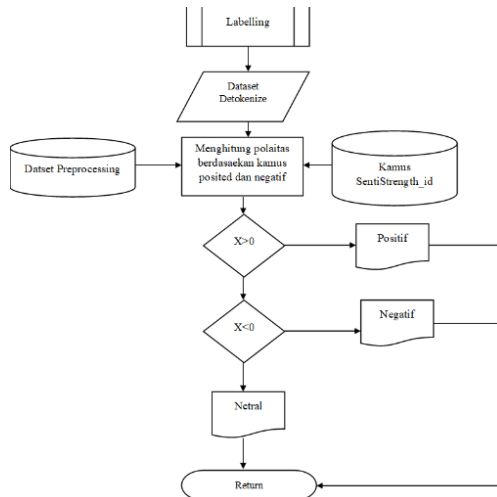
Nc: Banyaknya kemungkinan kata ke – c pada corpus

N : Banyaknya kata pada corpus

Jika suatu kata tidak ditemukan dalam korpus, maka akan dilakukan operasi edit distance dengan perubahan satu huruf. Jika kata tersebut masih tidak ditemukan, maka operasi edit distance dengan perubahan dua huruf akan diterapkan. Proses edit distance dengan perubahan dua huruf ini mirip dengan perubahan satu huruf, tetapi menggunakan kata hasil kemungkinan dari perubahan satu huruf. Jika kata tersebut tetap tidak ditemukan, maka kata tersebut akan dikembalikan seperti semula, meskipun terdeteksi salah. Namun, jika kata ditemukan dalam korpus, maka akan dilakukan pemeriksaan kembali terhadap semua kemungkinan kata dalam korpus dan dihitung peluang dari setiap kata yang ada. Rumus perhitungan probabilitas kata menurut Peter Norvig (Malik Iryana & Pandu Adikara, 2021) pada formula 1.

Labelling

Setelah melakukan preprocessing selanjutnya adalah proses labelling, menggunakan SentiStrength_id. Adapun alurnya adalah sebagai berikut.



Gambar 4. Labelling SentiStrength_id

If positive value > negative value then positive sentiment

If positive value < negative value then negative sentiment

If positive value = negative value then neutral sentiment

Pada proses labelling ini digunakan metode SentiStrength_id. Tahapan pada penerapan model ini diawali dengan menginput dataset yang sudah dilakukan pembersihan atau preprocessing, dilanjutkan dengan menginput enam kamus sentimen dari SentiStrength_id yaitu kamus sentimen, emosikon, idiom, boosterwords, negasi, dan kata tanya. Setelah itu akan dilakukan proses pengecekan pada setiap kata pada ulasan dengan kamus yang tersedia. Pemberian skor akan disesuaikan dengan kamus dan aturan ekstraksi fitur SentiStrength_id. Setelah diberi skor atau nilai, akan dilakukan pengambilan nilai untuk max positive dan max negative yang nantinya akan dibandingkan untuk menarik kesimpulan sentimen positif, atau negatif, atau netral. (Haryalesmana Wahid, 2016).

Modelling

Setelah dilakukan proses labelling dan pembobotan kata dengan TF-IDF, selanjutnya adalah proses pemodelan, dimana menggunakan SVM dengan kernel RBF dan metode multiclass One Against All (Suyanto, 2018):

Gaussian (Radial Basis Function (RBF))	$K(x, x_k) = \exp \left\{ \frac{- x_i x_j ^2}{2\sigma^2} \right\}$ <p>(1)</p>
---	---

HASIL DAN PEMBAHASAN

Pada proses awal dilakukan perolehan dataset dengan menggunakan web scrapping dilanjutkan dengan preprocessing. Adapun hasilnya sebagai berikut:

Tabel 1. Hasil Web Scrapping dan Preprocessing

Web Scrapping	Sangat mmbantu...dn respon ceoat
Preprocessing	Case Folding
	Cleaning
	Normalisasi
	Spelling Corrector
	Tokenizing
	Stopword
	Detokenizing

Pada proses spelling corrector terdapat beberapa kata salah eja yang tidak mengalami proses perbaikan sebagai contoh pada kata 'kalij' yang seharusnya 'kali'. Setelah dilakukannya proses di atas selanjutnya adalah proses labelling yaitu sebagai berikut:

Tabel 2. Proses Pembobotan SentiStrength_id

Labelling	['sangat', 'membantu', 'dan', 'respon', 'cepat']	-Kata 'sangat' pada kamus boosterword = 2
	['respon', 'cepat']	-Kata 'membantu' pada kamus sentiword = 4
		-Namun dikarenakan kata 'membantu' berada di depan kata 'sangat' maka polaritas berubah = 2 + 4 = 6.

Tabel 3. Hasil Sentimen

Labelling	Ulasan Dibobotkan	Max Positif	Max Negatif	Sentimen
	sangat membantu [6] respon cepat [4]	6	-1	Positif

Setelah dilakukan proses labelling dilanjutkan dengan TF-IDF dan dilanjutkan dengan pemodelan menggunakan SVM, yaitu sebagai berikut:

```

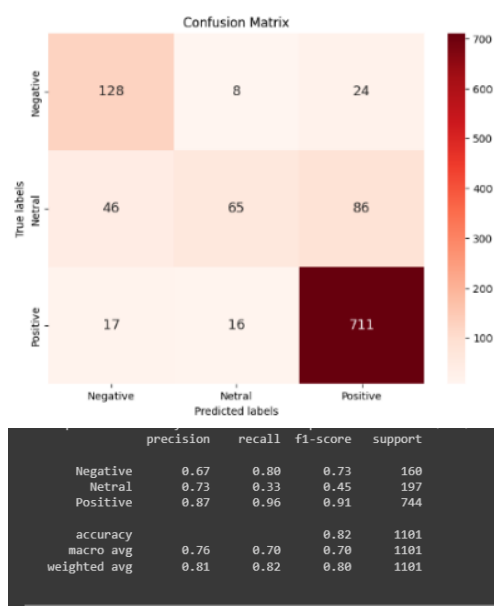
76] # Initialize the SVM classifier with OneVsRest approach
svm = SVC(C=10, kernel='rbf', gamma=1)
ovr_classifier = OneVsRestClassifier(svm)

# Fit the classifier on the training data
ovr_classifier = ovr_classifier.fit(X_train, y_train)

# Predict on the test data
test_predictions = ovr_classifier.predict(X_test)
test_predictions_df = pd.DataFrame(test_predictions, columns=['predict'])
  
```

Gambar 5. Modelling SVM

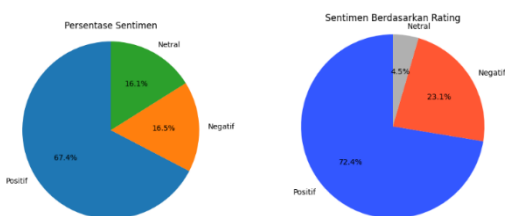
Setelah dilakukan pemodelan selanjutnya dilakukan evaluasi menggunakan confusion matrix yaitu sebagai berikut:



Gambar 6. Confusion Matrix dan Hasil

Selanjutnya setelah proses pelabelan adalah melakukan pemodelan. Setelah hasil prediksi dari pemodelan menggunakan Support Vector Machine (SVM) diperoleh, dilakukan evaluasi dengan menggunakan confusion matrix. Evaluasi dengan confusion matrix ini digunakan untuk mengetahui nilai akurasi pada analisis sentimen yang sudah dilakukan. Adapun diperoleh hasil setelah melakukan beberapa percobaan menggunakan pembagian split data, diperoleh hasil akurasi tertinggi dengan rasio 90:10 yaitu sebesar 82%.

Dari hasil proses analisis sentimen diperoleh juga hasil persentase sentimen menggunakan metode SentiStrength_id dan diperoleh hasil persentase jumlah pelabelan kelas yaitu 67,4% positif dengan 7.414 data, 16,5% negatif dengan 1.818 data, dan 16,1% netral dengan 1.772 data



Gambar 7. Hasil Persentase Pelabelan dan Berdasarkan Rating

Untuk membandingkan hasil terhadap sentimen sebelum dilabelkan, peneliti juga melakukan pengelompokan terhadap rating data ulasan dari Google Play Store sebelum dilakukan proses preprocessing dan labelling. Dimana untuk rating 5 dan 4 dilabelkan sebagai kelas positif, rating 3 sebagai netral, dan rating 2 dan 1 sebagai kelas negatif. Diperoleh hasil untuk kelas positif sebanyak 7.967 data, negatif 2.538 data, dan netral sebanyak 499 data dengan persentase untuk sentimen positif adalah 72,49%, negatif 23,1%, dan netral 4,5%. Dari persentase di atas dapat dilakukan perbandingan dengan hasil pelabelan dan diambil kesimpulan, bahwasannya persentase ketidaksesuaian pengguna dalam memberikan rating dengan ulasan diperoleh yaitu kelas positif sebesar 5% (553 data), kelas negatif 6,6% (720 data), dan kelas netral sebesar 11,6% (1.273 data).



Gambar 8. Word Cloud Positif dan Negatif



Gambar 9. Word Cloud Netral

Dari word cloud pada Gambar 8 bagian kiri, terdapat kata – kata yang mempengaruhi sentimen positif seperti ‘membantu’, ‘mudah’, ‘cepat’, dari hasil tersebut menunjukkan bahwa ulasan positif berpusat pada efisiensi aplikasi, kemudahan penggunaan, dan kualitas layanan yang diberikan oleh PLN melalui aplikasi PLN Mobile. Pada Gambar 8 bagian kanan mengandung kata – kata yang sering muncul pada sentimen negatif seperti ‘tidak’, ‘masalah’, ‘gangguan’, ‘lambat’, dari kata-kata tersebut menunjukkan bahwa ulasan negatif berpusat pada bagian masalah teknis, kegagalan fungsi, lambatnya aplikasi, serta ketidakpuasan respon dan pelayanan yang diberikan oleh aplikasi terkait layanan listrik. Pada Gambar 9 terdapat kata – kata yang sering muncul seperti kata ‘tidak’, ‘aplikasi’, ‘pln’, ‘token’, ‘listrik’, ‘pembayaran’, ‘lebih’, ‘beli’, ‘petugas’, ‘tolong’, ‘kenapa’, dan sebagainya. Kata – kata ini berdampak terhadap sentimen netral karna tidak memiliki pengaruh yang

signifikan terhadap sentimen positif dan negatif. Hal ini terkait pertanyaan mengenai aplikasi token dan listrik.

KESIMPULAN

Proses implementasi Peter Norvig sebagai spelling corrector terdapat pada tahapan preprocessing. Diawali dengan tahapan cleaning, case folding, normalisasi (slang word), spelling corrector, setelah itu dilanjutkan proses tokenizing, stopword, dan detokenize. Proses selanjutnya adalah labelling dengan SentiStrength_id menggunakan kamus-kamus yaitu sentimen, boosterwords, emoticon, idioms, negasi dan kata tanya yang sudah dibobotkan. Proses implementasi peter norvig dan sentistrength_id berhasil meningkatkan kualitas akurasi, dengan akurasi akhir sebesar 82% dengan rasio 90:10. Diperoleh dari hasil perbandingan kelas positif, netral, dan negatif terhadap 11.004 data hasil SentiStrength_id dengan ulasan berdasarkan rating, ketidaksesuaian pengguna memberikan rating dengan ulasannya adalah 5% untuk kelas positif dengan 553 data, 6,6%, untuk kelas negatif dengan 720 data, dan 11,6% untuk kelas netral dengan 1.273 data.

DISEMINASI

Artikel ini telah diseminasikan pada Seminar Nasional Teknologi Informasi dan Komunikasi (SEMNASITIK) APTIKOM Tahun 2024 yang diselenggarakan oleh Universitas Methodist Indonesia pada tanggal 24-26 Oktober 2024.

DAFTAR PUSTAKA

- Abdillah, W. F., Premana, A., & Herdian Bhakti, R. M. (2021). Analisis Sentimen Penanganan Covid-19 Dengan Support Vector Machine: Evaluasi Leksikon Dan Metode Ekstraksi Fitur. *Jurnal Ilmiah Intech : Information Technology Journal of UMUS*, 3(2), 160–170.
- Ahdiat, A. (2023, August 8). *67% Penduduk Indonesia Punya Handphone pada 2022, Ini Sebarannya*. DataBoks. <https://databoks.katadata.co.id/datapublish/2023/03/08/67-penduduk-indonesia-punya-handphone-pada-2022-ini-sebarannya>
- Alfred, R., & Teoh, R. W. (2019). Improving topical social media sentiment analysis by correcting unknown words automatically. *Communications in Computer and Information Science*, 937, 299–308. https://doi.org/10.1007/978-981-13-3441-2_23
- Annur, C. M. (2023, September 9). *Proporsi Penggunaan Internet Lewat Ponsel RI Lebih Tinggi dari Rerata Dunia*. DataBoks. <https://databoks.katadata.co.id/datapublish/2023/09/22/proporsi-penggunaan-internet-lewat-ponsel-ri-lebih-tinggi-dari-rerata-dunia>
- Asri, Y., Suliyanti, W. N., Kuswardani, D., & Fajri, M. (2022). Pelabelan Otomatis Lexicon Vader dan Klasifikasi Naive Bayes dalam menganalisis sentimen data ulasan PLN Mobile. *PETIR*, 15(2), 264–275. <https://doi.org/10.33322/petir.v15i2.1733>
- Aulia, B., Utomo, P. E. P., Khaira, U., & Suratno, T. (2021). Analisis Sentimen Tagar #IndonesiaTerseher Di Masa Covid-19 Menggunakan Metode Sentistrength. *Jurnal Komputer Dan Informatika*, 9(2), 207–213. <https://doi.org/10.35508/jicon.v9i2.4275>
- Ependi, U., Aliya, S., & Wibowo, A. (2023). Sentiment Analysis of Covid-19 Handling in Indonesia Based on Lexicon Weighting. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 12(1), 76–82. <https://doi.org/10.32736/sisfokom.v12i1.1615>
- Fatihah Rahmadayana, & Yuliant Sibaroni. (2021). Sentiment Analysis of Work from Home Activity using SVM with Randomized Search Optimization. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(5), 936–942. <https://doi.org/10.29207/resti.v5i5.3457>
- Fortune Indonesia 100. (2023). *Fortune Indonesia*. <https://www.fortuneidn.com/fortune-indonesia-100>
- Handayani, Y., Hakim, A. R., & Muljono. (2020). Sentiment analysis of Bank BNI user comments using the support vector machine method. *Proceedings - 2020 International Seminar on Application for Technology of Information and Communication: IT Challenges for Sustainability, Scalability, and Security in the Age of Digital Disruption, ISemantic 2020*, 202–207. <https://doi.org/10.1109/iSemantic50169.2020.9234230>
- Haryalesmana Wahid, D. (2016). Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity. *IJCCS*, 10(2), 207–218.
- Ilmawan, L. B., & Mude, M. A. (2020). Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store. *ILKOM Jurnal Ilmiah*, 12(2), 154–161. <https://doi.org/10.33096/ilkom.v12i2.597.154-161>
- Khaira, U., Johanda, R., Utomo, P. E. P., & Suratno, T. (2020). Sentiment Analysis Of Cyberbullying On Twitter Using SentiStrength. *Indonesian Journal of Artificial Intelligence and Data Mining*, 3(1), 21. <https://doi.org/10.24014/ijaidm.v3i1.9145>
- Kulkarni, A., & Shivananda, A. (2019). Natural language processing recipes: Unlocking text data with machine learning and deep learning

- using python. In *Natural Language Processing Recipes: Unlocking Text Data with Machine Learning and Deep Learning using Python*. Apress Media LLC. <https://doi.org/10.1007/978-1-4842-4267-4>
- Kusuma, A. T. A., & Ratnasari, C. I. (2023). Comparison Of Spell Correction In Bahasa Indonesia: Peter Norvig, LSTM, And N-Gram. *JIKO (Jurnal Informatika Dan Komputer)*, 6(3), 214–220. <https://doi.org/10.33387/jiko.v6i3.7072>
- Liu, B. (2010). *Sentiment Analysis and Subjectivity*.
- M Isa, S. (2017, October 4). *Sentiment Analysis Approaches and Methods*. Binus University Graduate Program . <https://mti.binus.ac.id/2017/10/04/1900/>
- Malik Iryana, T., & Pandu Adikara, P. (2021). Analisis Sentimen Masyarakat Terhadap Mass Rapid Transit Jakarta Menggunakan Metode Naïve Bayes Dengan Normalisasi Kata. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 5(6), 2548–2964. <http://j-ptiik.ub.ac.id>
- Martin, R., Santun Naga, D., & Christanti Mawardi, V. (2021). *Penggunaan Spelling Correction Dengan Metode Peter Norvig Dan N-Gram*.
- Mutammiah, Sujaini, H., & Nyoto, R. D. (2017). Analisis Perbandingan Metode Spelling Corrector Peter Norvig dan Spelling Checker BK-Trees pada Kata Berbahasa Indonesia. *Jurnal Sistem Dan Teknologi Informasi*, 5.
- Saefillah, C. A. (2023, April 6). *Fungsi & Tujuan Mobile Application Development Pada Bisnis*. Optify. <https://www.optify.co.id/website-development/fungsi-tujuan-mobile-application-development-pada-bisnis#:~:text=Pengembang%20aplikasi%20mobile%20bertujuan%20untuk,memberikan%20pengalaman%20pengguna%20yang%20optimal.>
- Sari, S., Utomo, P., Khaira, U., & Suratno, T. (2021). Analisis Sentimen Terhadap Komentar Beauty Shaming Di Media Sosial Twitter Menggunakan Algoritma SentiStrength. *IJRSE: Indonesian Journal of Informatic Research and Software Engineering*, 1, 71–78.
- Sasmoko, Prof. Dr. I. M. P. (2020, December 4). *Membangun Aplikasi Mobile Populer Oleh Memanfaatkan Umpan Balik Pengguna*. Binus University Research Interest Group. <https://research.binus.ac.id/edutech/2020/12/membangun-aplikasi-mobile-populer-oleh-memanfaatkan-umpan-balik-pengguna/>
- Sianipar, R., & Setiawan, E. B. (2015). *Pendeteksian Kekuatan Sentimen Pada Teks Tweet Berbahasa Indonesia Menggunakan Sentistrength*.
- Simanjuntak, M. S., Sujaini, H., & Safriadi, N. (2018). Spelling Corrector Bahasa Indonesia dengan Kombinasi Metode Peter Norvig dan N-Gram. *JEPIN (Jurnal Edukasi Dan Penelitian Informatika)*, 4.
- Suyanto. (2018). *Machine Learning Tingkat Dasar dan Lanjut, Edisi 2* (2nd ed.).
- Thomas, V. W. D., & Rumaisa, F. (2022). Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF. *Jurnal Media Informatika Budidarma*, 6(3), 1767. <https://doi.org/10.30865/mib.v6i3.4218>