

ANALISIS SENTIMEN MASYARAKAT TERHADAP ISU PEMALSUAN IJAZAH JOKO WIDODO DI MEDIA SOSIAL X MENGGUNAKAN METODE NAIVE BAYES

Teresa Martuah Purba✉, Nicolas Elsada Lahagu, Karolus Doweng Koten,
Richard Agung Orlando Berutu, Jumita Yohana Hutagalung

Fakultas Ilmu Komputer, Universitas Katolik Santo Thomas, Medan, Indonesia

Email: teresamartuahpurba26@gmail.com

DOI: <https://doi.org/10.46880/methoda.Vol15No3.pp248-254>

ABSTRACT

The alleged diploma forgery case involving Joko Widodo attracted great attention on social media, especially on the X platform. This research aims to analyze people's association with the issue by utilizing the Naive Bayes algorithm. The data used in this study consists of 1,000 tweets obtained through web crawling techniques based on certain keywords. The dataset is divided into two segments, 80% used for training data and 20% used for testing data. Of the 800 data used for training, 299 tweets (37.4%) showed positive sentiment, and 501 tweets (62.6%) showed negative sentiment. The remaining 200 tweets were used to test the model. The Naive Bayes model was then trained with the training data to identify existing sentiment patterns, and then tested with the test data to assess its classification accuracy. The findings of this study indicate that negative sentiment is superior to positive sentiment regarding the issue of Joko Widodo's diploma forgery. This reflects the public's tendency to respond to politically sensitive issues in the digital world. The Naive Bayes method proved to be quite powerful in classifying text-based public opinion with high accuracy, so it has the potential to become a tool in mapping public views on socio-political issues.

Keyword: Sentiment Analysis, Joko Widodo, Social Media Platform X, Naive Bayes.

ABSTRAK

Kasus dugaan pemalsuan ijazah yang melibatkan Joko Widodo menjadi sorotan besar di media sosial, terutama di platform X. Penelitian ini bertujuan untuk menganalisis perasaan masyarakat mengenai isu tersebut dengan memanfaatkan algoritma Naive Bayes. Data yang dipakai dalam penelitian ini terdiri dari 1.000 tweet yang diperoleh dengan metode web crawling menggunakan kata kunci tertentu. Data set terbagi menjadi dua segmen, yakni 80% untuk data pelatihan dan 20% untuk data pengujian. Dari 800 data yang dipakai untuk pelatihan, diperoleh hasil klasifikasi awal dengan 299 tweet (37,4%) yang memiliki sentimen positif, serta 501 tweet (62,6%) yang memiliki sentimen negatif. Sisa data sebanyak 200 tweet digunakan untuk menguji model. Model Naive Bayes selanjutnya dilatih dengan data pelatihan untuk mengenali pola-pola sentimen yang ada, lalu diuji dengan data pengujian untuk menilai akurasi klasifikasi. Temuan dari penelitian ini mengindikasikan bahwa sentimen negatif lebih unggul dibandingkan sentimen positif terkait dengan isu pemalsuan ijazah Joko Widodo. Ini menunjukkan kecenderungan masyarakat dalam menanggapi isu sensitif secara politik di ranah digital. Metode Naive Bayes terbukti efektif dalam mengklasifikasikan opini publik dari teks dengan akurasi yang tinggi, sehingga dapat berfungsi sebagai alat dalam memetakan pendapat masyarakat mengenai isu-isu sosial-politik.

Kata Kunci: Sentimen Analisis, Joko Widodo, Platform Sosial Media X, Naive Bayes.

PENDAHULUAN

Media sosial telah menjadi ruang publik digital yang memungkinkan masyarakat untuk menyampaikan opini, emosi, dan persepsi terhadap berbagai isu sosial, politik, dan budaya secara bebas dan terbuka. Salah satu isu yang sempat mengemuka dan menimbulkan kontroversi adalah dugaan pemalsuan ijazah oleh Presiden Republik Indonesia, Joko Widodo. Isu ini menjadi viral di berbagai platform media sosial, khususnya platform X (sebelumnya Twitter), yang dikenal sebagai wadah utama diskusi daring masyarakat Indonesia terhadap isu-isu nasional.

Respons masyarakat terhadap isu tersebut sangat beragam, mulai dari dukungan hingga penolakan, yang sering kali diungkapkan dalam bentuk teks bernuansa emosional, baik positif maupun negatif. Oleh karena itu, perlu dilakukan analisis lebih lanjut untuk memahami kecenderungan sentimen publik terhadap isu ini secara sistematis. Salah satu pendekatan yang umum digunakan untuk tujuan ini adalah analisis sentimen berbasis teks, yang bertujuan mengkategorikan opini pengguna ke dalam kelas sentimen tertentu, seperti positif, negatif, atau netral.

Dalam penelitian ini, digunakan metode *Naive Bayes* yang merupakan salah satu algoritma pembelajaran mesin berbasis probabilitas yang efektif dalam tugas klasifikasi teks, termasuk analisis sentimen. Metode ini mampu mengidentifikasi pola bahasa dan kata kunci yang berhubungan dengan sentimen tertentu secara cepat dan efisien, bahkan dengan jumlah data yang besar.

Penelitian ini tidak hanya penting untuk mengetahui persepsi masyarakat terhadap Presiden Joko Widodo dalam konteks tuduhan pemalsuan ijazah, tetapi juga relevan dalam memahami bagaimana opini publik terbentuk dan tersebar di ruang digital. Selain itu, hasil analisis ini dapat menjadi masukan bagi para pembuat kebijakan, peneliti sosial-politik, serta masyarakat umum dalam menyikapi isu-isu kontroversial secara lebih bijak.

KAJIAN LITERATUR

Media Sosial X sebagai Representasi Opini Publik

Perkembangan media sosial telah mengubah cara masyarakat menanggapi isu publik. Platform X (dulu Twitter) menyediakan ruang bagi pengguna untuk menyampaikan pendapat terhadap isu-isu aktual secara langsung dan terbuka. Media sosial memungkinkan pengumpulan data dalam skala besar dan real-time, menjadikannya alat yang ideal untuk analisis opini publik. Menurut (Naraswati et al., 2021) media sosial seperti Twitter dapat merepresentasikan opini masyarakat secara luas karena tweet mengandung ekspresi emosional yang mencerminkan sikap pengguna terhadap suatu isu.

Analisis Sentimen dan Relevansinya dalam Penelitian Sosial⁷

Analisis sentimen adalah proses identifikasi dan ekstraksi opini dari teks untuk menentukan sikap atau sentimen (positif, negatif, atau netral) terhadap subjek tertentu. Dalam kajian ilmu komputer dan linguistik, pendekatan ini digunakan untuk memahami bagaimana suatu entitas atau isu dipersepsikan oleh masyarakat. (Prasetyo Wibowo et al., 2023) menekankan bahwa analisis sentimen penting digunakan dalam konteks sosial-politik karena memberikan gambaran seberapa besar kepercayaan atau ketidakpuasan masyarakat terhadap pemerintah dan tokoh politik.

Twitter merupakan platform media sosial yang sangat efektif dalam menggambarkan persepsi publik terhadap isu-isu aktual, termasuk layanan digital dan kebijakan publik (Saputra et al., 2024; Siringoringo et al., 2023). Mereka menggunakan algoritma *Naive Bayes* untuk mengkaji opini masyarakat, dan hasilnya menunjukkan bahwa media sosial dapat mencerminkan dinamika sosial secara cepat.

Naive Bayes dalam Analisis Sentimen

Metode *Naive Bayes* adalah algoritma pembelajaran mesin berbasis probabilitas yang populer dalam klasifikasi teks (Pakpahan et al., 2024). Metode ini bekerja dengan menghitung probabilitas setiap kelas (positif, negatif)

terhadap kemunculan kata-kata tertentu dalam teks. Dalam penelitian (Ramadhani et al., 2021), algoritma ini digunakan untuk mengklasifikasikan opini masyarakat terhadap COVID-19 dan menunjukkan tingkat akurasi sebesar 87%. Sementara itu, (Bagus Sasmita et al., 2022) menunjukkan bahwa dengan *preprocessing* yang baik (penghapusan *stopword*, *stemming*, *tokenizing*), akurasi klasifikasi Naive Bayes dalam menganalisis sentimen terhadap kebijakan PPKM di media sosial dapat mencapai 77,2%.

Penelitian Terdahulu yang Relevan

Penelitian sebelumnya menunjukkan bahwa isu-isu sensitif seperti politik, bencana, atau kebijakan publik cenderung memunculkan reaksi negatif di media sosial. (Maksun et al., 2021) mengkaji isu bencana di Kalimantan Selatan dengan metode Naive Bayes, dan mendapati bahwa sentimen negatif mendominasi respons masyarakat.

(Asthakhuroh et al., 2023) menerapkan metode Naive Bayes untuk menganalisis opini masyarakat terhadap program pemerintah, dan membuktikan bahwa metode ini mampu memberikan akurasi tinggi dengan efisiensi proses komputasi yang baik.

(Susanto et al., 2021) menguatkan hal tersebut dengan menyatakan bahwa *Naive Bayes* efektif untuk data teks dalam bahasa Indonesia dan dapat digunakan untuk klasifikasi sentimen yang bersifat cepat dan ringan secara proses.

METODE PENELITIAN

Metodologi dalam penelitian ini merujuk pada alur proses analisis sentimen. Tahapan-tahapan tersebut meliputi pengumpulan data, *preprocessing*, clustering, labeling, klasifikasi dengan algoritma *Naive Bayes*, hingga evaluasi model. Berikut adalah penjelasan rinci tiap tahapan:

Pengumpulan Data (*Training Data*)

Langkah awal dalam penelitian ini adalah mengumpulkan data latih (*training data*) dari media sosial X (dahulu Twitter). Data yang dikumpulkan berupa tweet dari pengguna yang berdomisili di Indonesia dan mengandung opini

atau pernyataan mengenai isu dugaan pemalsuan ijazah Presiden Joko Widodo. Tweet diperoleh dengan menggunakan kata kunci seperti “ijazah Jokowi”, “ijazah palsu”, dan “UGM” agar relevan dengan isu yang dibahas.

Dari hasil *crawling*, terkumpul 1.000 tweet yang kemudian dibagi menjadi:

- 800 tweet (80%) digunakan sebagai data pelatihan (*training data*), dan
- 200 tweet (20%) digunakan sebagai data pengujian (*testing data*).

Preprocessing (Pelatihan dan Pengujian)

Tweet yang diperoleh tidak bisa langsung digunakan karena masih mengandung banyak unsur yang tidak relevan, seperti karakter khusus, tautan, mention (@), dan kata-kata tidak bermakna (*stopwords*). Oleh karena itu, dilakukan tahapan *preprocessing* untuk membersihkan dan menyederhanakan data agar lebih siap digunakan dalam proses analisis.

Beberapa langkah dalam *preprocessing* ini meliputi:

- *Case folding* – mengubah seluruh teks menjadi huruf kecil.
- *Tokenizing* – memisahkan kalimat menjadi kata-kata individual.
- *Stopword removal* – menghapus kata-kata umum yang tidak membawa makna penting.
- *Stemming* – mengembalikan kata ke bentuk dasarnya.
- Pembersihan karakter khusus – menghapus tanda baca, emoji, dan tautan yang tidak dibutuhkan.

Tabel 1. Contoh Hasil *Preprocessing*

No	Tweet Mentah	Hasil <i>Preprocessing</i>
1	"Ijazah Jokowi itu asli kok, sudah diklarifikasi pihak UGM! 👍 #JokowiAsli"	ijazah jokowi itu asli kok sudah diklarifikasi pihak ugm jokowiasli
2	"Link ini kasih bukti bahwa ijazah jokowi palsu: http://bit.ly/xyz123 #palsu"	link ini kasih bukti bahwa ijazah jokowi palsu

Clustering dan Labeling

Setelah data dibersihkan, langkah berikutnya adalah mengelompokkan tweet berdasarkan kemiripan kata dan konteks melalui proses *clustering*. Tahap ini bertujuan untuk menyederhanakan proses pelabelan. Kemudian, dilakukan *labeling* secara manual untuk menentukan kategori sentimen setiap tweet, yaitu:

- Positif: jika tweet berisi dukungan atau pembelaan terhadap Presiden Joko Widodo.
- Negatif: jika tweet bersifat menuduh atau menyudutkan terkait isu ijazah.

Tabel 2. Contoh Proses *Labeling*

No	Tweet	Label
1	Jangan percaya hoaks, ijazah presiden kita sudah resmi	Positif
2	Ijazah Jokowi palsu, pemerintah menutup-nutupinya	Negatif

Pengujian Data (*Testing Data*)

Data pengujian terdiri atas 200 tweet yang telah melewati tahapan preprocessing. Data ini tidak digunakan dalam proses pelatihan agar hasil evaluasi terhadap model menjadi lebih objektif. Tujuannya adalah untuk melihat seberapa baik model klasifikasi dapat bekerja pada data yang benar-benar baru.

Klasifikasi (*Classification*)

Setelah proses pelabelan selesai, dilakukan klasifikasi menggunakan algoritma Naive Bayes. Algoritma ini termasuk dalam metode *probabilistic classification* yang bekerja berdasarkan *Teorema Bayes* dengan asumsi bahwa setiap fitur (kata dalam teks) bersifat independen.

Walaupun asumsi independen ini tidak selalu berlaku dalam bahasa alami, *Naive Bayes* tetap menjadi pilihan yang populer karena hasilnya yang cukup akurat dan efisien untuk data berukuran besar. Dalam konteks ini, algoritma digunakan untuk mempelajari pola sentimen dari data pelatihan, lalu memprediksi sentimen dari data pengujian.

Jenis penelitian ini tergolong sebagai kuantitatif deskriptif, karena bertujuan untuk menguji kebenaran hipotesis berdasarkan data statistik dari opini masyarakat dalam bentuk teks (Mutiar, 2020).

Kelas Sentimen (*Sentiment Class*)

Hasil dari klasifikasi akan menghasilkan kelas sentimen, yang terbagi menjadi:

- Sentimen positif, dan
- Sentimen negatif.

Klasifikasi ini merepresentasikan sikap publik terhadap isu pemalsuan ijazah Jokowi dan menjadi dasar untuk menarik kesimpulan mengenai persepsi publik di media sosial X terhadap isu yang tengah berkembang.

Evaluasi Model

Langkah terakhir adalah mengevaluasi performa model klasifikasi yang telah dibangun. Evaluasi ini dilakukan dengan membandingkan hasil prediksi model terhadap label asli dari data pengujian. Metrik yang digunakan adalah *Confusion Matrix* untuk mengetahui distribusi klasifikasi benar dan salah.

Melalui evaluasi ini, dapat diketahui sejauh mana algoritma *Naive Bayes* efektif dalam mengenali opini publik terhadap isu sosial-politik, khususnya terkait dugaan pemalsuan ijazah Joko Widodo di media sosial X.

HASIL DAN PEMBAHASAN

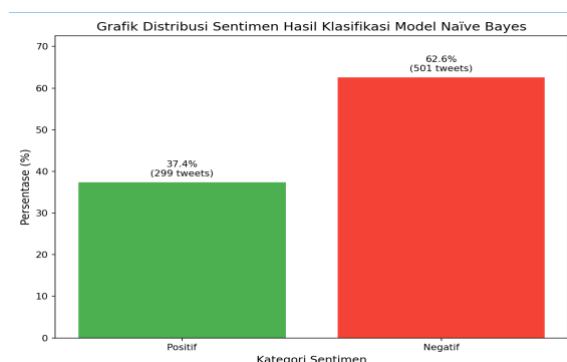
Dengan menggunakan metode *Naive Bayes*, analisis ini dilakukan untuk mempelajari sentimen masyarakat terhadap masalah pemalsuan ijazah Joko Widodo di platform X. Data yang dianalisis terdiri dari 1.000 tweet berbahasa Indonesia yang dikumpulkan melalui teknik scraping web dari 01 Mei 2025 hingga 23 Mei 2025. Kata kunci yang digunakan meliputi “pemalsuan ijazah Joko Widodo”, “ijazah palsu Jokowi”, “kontroversi ijazah Jokowi”, dan istilah terkait lainnya. Menggunakan algoritma *Naive Bayes* yang digunakan oleh perangkat lunak *RapidMiner*, tweet-tweet ini dikategorikan ke dalam dua kategori sentimen: positif dan negatif.

Dari 1.000 tweet, 80% (800 tweet) digunakan sebagai data latih untuk melatih model klasifikasi, dan 20% (200 tweet) digunakan sebagai data uji untuk mengevaluasi performa model. Hasil klasifikasi pada data latih menunjukkan bahwa 501 tweet (62,6%) termasuk dalam kategori sentimen negatif, dan 299 tweet (37,4%) termasuk dalam kategori sentimen positif. Distribusi sentimen ini ditampilkan dalam tabel berikut:

Tabel 3. Distribusi Klasifikasi Sentimen pada Data Latih

Sentimen	Jumlah Tweet	Persentase
Positif	299	37,4%
Negatif	501	62,6%
Total	800	100%

Distribusi ini juga divisualisasikan dalam bentuk diagram batang untuk menggambarkan kecenderungan opini masyarakat secara lebih jelas.



Gambar 1. Grafik Distribusi Sentimen Hasil Klasifikasi Model *Naive Bayes*

Tingginya proporsi sentimen negatif (62,6%) menunjukkan bahwa mayoritas pengguna X menyatukan pandangan kritis terhadap isu pemalsuan ijazah Joko Widodo. Banyak tweet negatif mencerminkan ketidakpercayaan, kecurigaan, atau penolakan terhadap kredibilitas Joko Widodo terkait isu ini. Sebagai contoh, beberapa tweet negatif yang diidentifikasi menyatakan:

- “Ijazah palsu Jokowi jadi bukti kalau integritas pemimpin kita patut dipertanyakan. Rakyat pantas tahu kebenaran!”

- “Gimana mau percaya sama pemimpin kalau urusan ijazah aja gak jelas? Ini serius loh!”

Tweet-tweet ini menunjukkan adanya kekecewaan dan keraguan masyarakat terhadap keabsahan ijazah Joko Widodo, yang dianggap mencerminkan masalah integritas dan transparansi dalam kepemimpinan. Sentimen negatif ini juga mengindikasikan bahwa isu ini memicu diskusi yang intens di platform X, dengan banyak pengguna yang menuntut klarifikasi atau kebenaran terkait tuduhan tersebut.

Di sisi lain, 37,4% tweet yang termasuk dalam kategori positif menunjukkan adanya sebagian masyarakat yang mendukung atau mempercayai keabsahan ijazah Joko Widodo, atau mengangap isu ini sebagai tuduhan yang tidak berdasar. Tweet-tweet positif ini cenderung membela Joko Widodo atau menolak narasi pemalsuan ijazah sebagai fitnah. Beberapa contoh tweet positif adalah:

- “Isu ijazah Jokowi cuma fitnah politik. Dia udah buktikan kerja nyata, ngapain ributin ijazah?”
- “Orang-orang yang nyebar isu ijazah palsu cuma mau jatuhin Jokowi. Gak ada bukti kuat, cuma omong kosong.”
- “Jokowi udah terbukti memimpin dengan baik. Ijazah itu cuma alasan buat nyerang dia.”

Tweet-tweet ini menunjukkan sebagian pengguna X percaya bahwa masalah pemalsuan ijazah adalah serangan politik yang tidak memiliki dasar kuat, seperti yang ditunjukkan dalam tweet-tweet ini. Mereka lebih fokus pada tindakan Joko Widodo selama kepemimpinannya. Sebagian besar dukungan didasarkan pada gagasan bahwa tuduhan tersebut merupakan upaya untuk merusak reputasinya.

Hasil analisis ini menunjukkan bahwa masalah pemalsuan ijazah Joko Widodo menyebabkan polarisasi opini di kalangan pengguna X. Sentimen negatif mendominasi, menunjukkan bahwa masyarakat cenderung skeptis dan kritis terhadap masalah ini, yang dapat memengaruhi kepercayaan publik terhadap orang-orang publik yang relevan.

Namun, sentimen positif menunjukkan bahwa masih ada bagian masyarakat yang percaya pada Joko Widodo, terutama mereka yang melihat tuduhan ini sebagai taktik politik.

Hasil ini menunjukkan bahwa dalam menangani masalah sensitif seperti ini, komunikasi publik yang jujur dan berdasarkan fakta sangat penting. Untuk mengurangi spekulasi dan meningkatkan kepercayaan masyarakat, pemerintah atau pihak terkait harus merespons dengan cepat dan jelas. Selain itu, temuan ini menunjukkan bahwa media sosial, khususnya X, memainkan peran penting dalam demokrasi digital, di mana pendapat publik diperdebatkan dan dibentuk secara terbuka. Karena itu, sangat penting bagi pihak yang terlibat dalam masalah ini untuk memanfaatkan platform ini secara strategis untuk melawan narasi yang tidak benar, membangun komunikasi yang mendidik, dan mendukung literasi politik masyarakat.

KESIMPULAN

Tujuan penelitian ini adalah untuk menggunakan metode Naive Bayes untuk mengevaluasi reaksi masyarakat pengguna platform X terhadap masalah pemalsuan ijazah Joko Widodo. Hasil analisis terhadap 1.000 tweet berbahasa Indonesia yang dikumpulkan melalui web scraping menunjukkan bahwa pada data latih (80%, 800 tweet), 62,6% (501 tweet) termasuk dalam kategori sentimen negatif, dan 37,4% (299 tweet) termasuk dalam kategori sentimen positif. Hasil pengujian pada data uji (20%, 200 tweet, dan 200 tweet, masing-masing

Menurut hasil analisis sentimen, sebagian besar masyarakat di platform X menanggapi masalah pemalsuan ijazah Joko Widodo dengan sentimen negatif, yang menunjukkan skeptisisme dan ketidakpercayaan terhadap informasi yang beredar. Banyak pengguna menganggap masalah ini sebagai agenda politik atau fitnah yang disebarluaskan untuk kepentingan tertentu. Sentimen negatif ini menunjukkan masalah dalam komunikasi politik digital, di mana ketidakpercayaan publik dapat meningkat karena kurangnya informasi resmi yang jelas.

DAFTAR PUSTAKA

- Asthakhuroh, A., Komarudin, R., & Kholifah, D. N. (2023). Sentiment Analysis With A Case Study Of Practice Card On Twitter Social Media Using Naive Bayes Method. *Jurnal Techno Nusa Mandiri*, 20(1), 8–13. <https://doi.org/10.33480/techno.v20i1.3709>
- Bagus Sasmita, A., Rahayudi, B., & Muflikhah, L. (2022). *Analisis Sentimen Komentar pada Media Sosial Twitter tentang PPKM Covid-19 di Indonesia dengan Metode Naive Bayes* (Vol. 6, Issue 3). <http://j-ptiik.ub.ac.id>
- Maksun, A. M., Sari, Y. A., & Rahayudi, B. (2021). *Analisis Sentimen pada Twitter Bencana Alam di Kalimantan Selatan menggunakan Metode Naive Bayes* (Vol. 5, Issue 12). <http://j-ptiik.ub.ac.id>
- Mutiara, E. (2020). Algoritma Klasifikasi Naive Bayes Berbasis Particle Swarm Optimization Untuk Prediksi Penyakit Tuberculosis (Tb). *Jurnal SwabumI*, 8(1), 46–58.
- Naraswati, N. P. G., Nooraeni, R., Rosmilda, D. C., Desinta, D., Khairi, F., & Damaiyanti, R. (2021). Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification. *Sistemasi*, 10(1), 222-238.
- Pakpahan, I. S., Dumayanty, I. S., & Manurung, S. (2024). Sistem Pakar Deteksi Penyakit Tanaman Cabai Metode Naive Bayes. *METHOTIKA: Jurnal Ilmiah Teknik Informatika*, 4(2), 10-20.
- Prasetyo Wibowo, M., Amini, S., & Kusumaningsih, D. (2023). *2 nd Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI) 21 Maret 2023- Jakarta* (Vol. 2, Issue 1).
- Ramadhani, T., Sari, Y. A., & Santoso, E. (2021). Analisis sentimen masyarakat indonesia terhadap covid-19 pada media sosial twitter menggunakan metode naive bayes. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 5(12), 5680-5686.
- Saputra, A. D., Prasetyo Budiman, D., Reynanda, R. M., & Sari, A. P. (2024). Analisis sentimen aplikasi gojek pada twitter menggunakan algoritma naive bayes. *JAMASTIKA: Jurnal Mahasiswa Teknik Informatika*, 3(2), 107-116
- Siringoringo, R., Jamaluddin, J., Perangin-angin, R., Harianja, E. J. G.,

Lumbantoruan, G., & Purba, E. N. (2023).
Model bidirectional lstm untuk
pemrosesan sekuensial data teks spam.
*METHOMIKA: Jurnal Manajemen
Informatika & Komputerisasi Akuntansi*,
7(2), 265-271.

Susanto, A., Maula, M. A., Mulyono, I. U. W.,
& Sarker, M. K. (2021). Sentiment
Analysis on Indonesia Twitter Data Using
Naïve Bayes and K-Means Method.
Journal of Applied Intelligent System,
6(1), 40–45.
<https://doi.org/10.33633/jais.v6i1.4465>