

ANALISIS SENTIMEN PRODUK HERBAL JAMU PADA MEDIA SOSIAL INSTAGRAM

Daniel Kurniawan¹, HD Purnomo², Ade Iriani³

^{1,2,3} Program Studi Magister Sistem Informasi, Fakultas Teknologi Informasi,
Universitas Kristen Satya Wacana

e-mail: 972020016@student.uksw.edu

Abstrak

Perubahan gaya hidup sehat dan peningkatan konsumsi tanaman obat dalam masyarakat di Indonesia memberikan dampak positif seperti peningkatan kapasitas dan jumlah industri serta pelaku usaha kecil di bidang herbal jamu. Produk herbal jamu bagi sebagian besar masyarakat Indonesia telah dipercaya dan lebih aman dalam mengobati penyakit. Produk herbal jamu biasa dimanfaatkan untuk merawat pasien penderita kelompok penyakit demam, hipertensi, luka, kolesterol, maag, malaria, dan infeksi saluran kencing. Tentunya pemilihan tanaman obat untuk merawat suatu penyakit juga dipengaruhi oleh budaya atau kepercayaan wilayah tertentu. Dalam kegiatan pemasarannya, produk herbal jamu juga memanfaatkan teknologi yang telah berkembang seperti penggunaan media sosial yang memiliki potensi untuk mendapatkan jangkauan konsumen dan pemasaran yang lebih luas hingga peningkatan omset dari penjualan produk. Media sosial yang dirasa mudah untuk digunakan pelaku usaha skala kecil maupun besar dalam pemasaran produk herbal adalah Facebook dan Instagram. Penelitian ini bertujuan untuk mengungkap sentimen pengguna media sosial Instagram terhadap tren penjualan herbal jamu. Hasil penelitian akan mengungkap signifikansi penggunaan media sosial Instagram pada pemasaran produk herbal jamu.

Kata kunci: Analisis Sentimen, klasifikasi Naïve Bayes, Analisis Jalur, Produk Herbal

Abstract

The changes in healthy lifestyles and increased consumption of medicinal plants in Indonesian society have had positive impacts, such as increasing the capacity and number of industries and small business actors in the herbal medicine sector. Herbal products for most Indonesians people have been trusted and are safer for treating diseases. Herbal products usually treat patients with fever, hypertension, wounds, cholesterol, ulcers, malaria, and urinary tract infections. Of course, the selection of medicinal plants to treat a disease is also influenced by the culture or beliefs of a particular region. In its marketing activities, herbal products also utilize technology that has developed, such as social media, which has the potential to reach more comprehensive consumers and marketing reach to increase turnover from product sales. Social media that are easy for small and large-scale businesses to use in marketing herbal products are Facebook and Instagram. This study aims to reveal the sentiments of Instagram social media users towards the selling trend of herbal medicine. The study's results will reveal the significance of using Instagram social media in marketing herbal products.

Keyword: Sentiment Analysis, Naïve Bayes Classification, Path Analysis, Herbal Product

PENDAHULUAN

Analisis sentimen atau penambangan opini merupakan teknik pemrosesan bahasa sehari-hari yang dipakai untuk mengekstrak atau memperoleh nilai perasaan yang terkandung dalam sebuah teks [1]. Sentimen analisis studi komputasi mengenai opini, sentimen, emosi, *attitude* seseorang terhadap sebuah produk, layanan, peristiwa, topik dan atribut lainnya [2]. Teknik seperti ini dipergunakan juga dalam meningkatkan kualitas

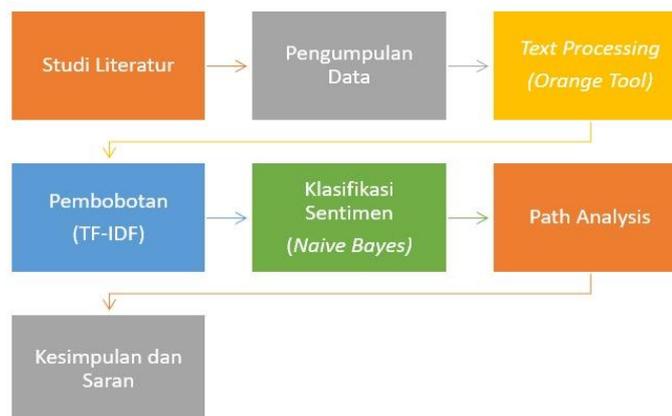
pendidikan seperti dalam penentuan keputusan, konsep pedagogik, dan sistem evaluasi [3]. Mirip dengan analisa sentimen, elektronik *word of mouth* yang dipengaruhi oleh opini orang juga dapat memberikan pengaruh seperti pada kriptokurensi, reputasi, penjualan, dan peringkat saham yang muncul dari pendapat media sosial seperti *twitter* [4]. Untuk contohnya dalam memprediksi pergerakan nilai *stock* harga saham yang dibarengi dengan analisa teks akan lebih cocok dipergunakan untuk mendukung sebuah keputusan daripada untuk mengambil keputusan. Penilaian manusia tetap menjadi penting saat teks yang dianalisa bersifat ambigu atau memiliki makna yang kurang jelas [4].

Meningkatnya industri dan usaha kecil yang bergerak pada bidang herbal jamu di Indonesia mengindikasikan adanya peningkatan dan perubahan gaya hidup masyarakat untuk mengkonsumsi tanaman obat [5]. Jamu atau obat tradisional bagi sebagian masyarakat telah dipercaya dan lebih aman dalam mengobati penyakit. Obat tradisional atau jamu juga memberikan manfaat bagi perawatan penderita batu saluran kemih dan diabetes melitus [6][7]. Meski begitu tetap terdapat faktor lain seperti budaya dan kepercayaan dalam memilih obat herbal [8]. Pemasaran sebuah produk herbal jamu dengan memanfaatkan media sosial kini menjadi sebuah kebutuhan dan dapat memperluas jangkauan pemasaran bahkan meningkatkan omset penjualan produk [9][10].

Pada sistem pemasaran digital, komunikasi antara pembeli dengan penjual merupakan suatu prioritas dan dapat membantu dalam peningkatan kualitas dan layanan penjualan. Keterlibatan pembeli memberikan respon kepada penjual umumnya dimudahkan dengan memberikan pendapat pada sebuah kolom di media social secara khusus *Instagram* yang dimiliki oleh *Facebook*. Pendapat konsumen akan bermanfaat seperti untuk sentimen analisis yang dipakai untuk mengklasifikasi pendapat atau ulasan menjadi pendapat positif atau negatif untuk membantu pengembangan dan peningkatan pemasaran produk [12]. Dengan begitu tidak hanya penjual dapat melakukan pembenahan dalam layanan, tetapi konsumen juga akan memiliki peran menilai dan berpartisipasi mengembangkan produk herbal jamu yang berkualitas baik.

METODE PENELITIAN

Penelitian ini melakukan analisis sentimen menggunakan metode pemrosesan dan pembobotan teks, klasifikasi *Naïve Bayes* dan Analisis Jalur untuk memperoleh hasil. Bagan yang menunjukkan alur penelitian ditampilkan sebagai berikut:



Gambar 1. Alur Penelitian Analisis Sentimen Produk Herbal

1.1 Studi literatur

Studi literatur akan digunakan apabila topik penelitian telah ditentukan dengan melakukan pencarian referensi, teori, dan temuan-temuan dari penelitian yang terkait dengan penelitian. Dalam penelitian ini, studi literatur digunakan untuk membantu membentuk perspektif dan mengumpulkan data yang diperlukan untuk melakukan analisis sentimen.

1.2 Pengumpulan data

Pengumpulan data dalam penelitian ini menggunakan metode *web scraping* atau teknik pengumpulan data pada halaman situs. Situs yang menjadi sasaran penelitian ini adalah *Instagram* dan data tren dari *Google trends* mengenai produk Herbal Sido Muncul periode 2019 hingga 2023. Data yang dikumpulkan berasal dari komentar data setiap konten yang diposting oleh akun Instagram @sidomunculstore dari Januari 2019 hingga Januari 2023. Data yang diambil dari situs Instagram adalah elemen situs web yang menunjukkan akun pembuat komentar, komentar, dan tanggal komentar ditulis.

1.3 Pemrosesan teks

Kumpulan teks yang telah terkumpul selanjutnya akan dipecah menjadi elemen-elemen yang lebih kritis sesuai dengan kebutuhan penelitian ini. Langkah-langkah yang dilakukan dalam melakukan pemrosesan teks adalah sebagai berikut:

- a. *Case Folding*, dilakukan dengan mengubah teks menjadi huruf kecil sehingga memiliki format case yang seragam.
- b. *Number Removal*, langkah ini dilakukan dengan menghapus nomor di setiap komentar.
- c. *Punctuation Removal*, langkah ini dilakukan untuk menghilangkan tanda baca.
- d. *Whitespace Removal*, langkah ini dilakukan untuk menghilangkan karakter tidak bernilai/kosong.
- e. *Tokenization*, tahap ini dilakukan untuk memecah kalimat sehingga hasil ini mengubah kalimat menjadi kumpulan array dari beberapa kata.
- f. *Stopword removal*, dan *Stemming*, tahap ini merupakan langkah menghilangkan konjungsi yang kurang berhubungan dengan sentimen.

1.4 Pembobotan Kata

Tahapan pembobotan kata diberikan untuk memberikan nilai pada kata dengan menghitung setiap kata pada dataset [13]. Pembobotan dilakukan setelah melakukan tahapan pemrosesan teks seperti formalisasi, tokenisasi, *stop removal*, dan *stem*, serta dapat dilihat pentingnya setiap kata dalam dataset. Kemudian proses klasifikasi dapat dilakukan setelahnya. Pembobotan kata menggunakan metode TF-IDF yang mudah dipelajari dan diterapkan [14].

1.5 Klasifikasi Naïve Bayes

Metode *Naïve bayes* merupakan salah satu metode digunakan untuk melakukan analisis sentimen terhadap data berupa teks seperti melakukan klasifikasi ulasan pada situs *tripAdvisor* dan *Google Review* pada produk kuliner bakso malang [13], mengetahui persepsi masyarakat mengenai transportasi online [15] dan aplikasi dompet elektronik pada media sosial *twitter* [16]. Untuk melakukan prediksi, kriteria evaluasi yang

diperlukan misalnya *accuracy*, *f1 score*, *recall*, *precision*, atau dengan menggunakan *confusion matrix* [17]. *Accuracy* dipakai untuk mengukur nilai prediksi pada tabel atau kelas dengan *score* paling tinggi 1 dan *score* paling rendah adalah 0. *Recall* merupakan hasil dari nilai True Positif dibagi dengan jumlah nilai True Positif dan False Negatif. *Precision* merupakan hasil nilai True Positif dibagi dengan jumlah nilai True Positif dan False Positif. Sedangkan *F1 score* yakni nilai keseimbangan antara *recall* dan *precision* [17].

1.6 Path Analysis

Path Analysis merupakan analisis statistik yang dipergunakan untuk menganalisis sebuah relasi antara sebab dan akibat baik itu secara langsung maupun tidak langsung terhadap satu atau lebih variabel yang dipakai untuk menemukan suatu hipotesis [18]. Dalam penelitian ini X1 merupakan teks positif, X2 teks negative, X3 teks netral dan Y adalah tren produk herba.

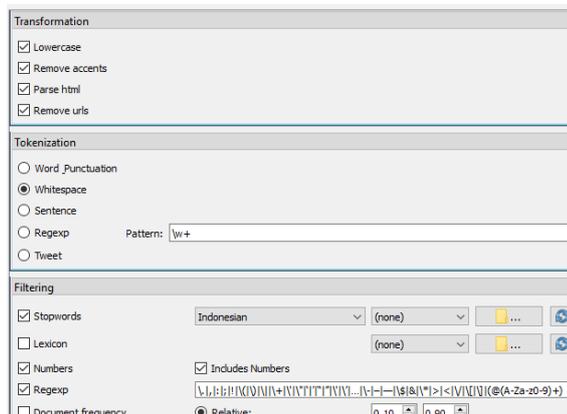
Setelah melewati alur di atas, hasil penelitian akan dilakukan pembahasan untuk ditarik kesimpulan dan saran sehingga dapat memberikan sumbangsih bagi keilmuan sistem informasi di masa depan.

HASIL DAN PEMBAHASAN

1.7 Dataset

Dataset diambil menggunakan metode *web scraping* dan memperoleh hasil data berjumlah 1906 komentar yang belum terklasifikasi. Dataset kemudian dibagi menjadi 2 bagian dengan perbandingan 20:80 yaitu 20% keseluruhan data untuk data training dan 80% keseluruhan data untuk data testing.

1.8 Pemrosesan teks



Gambar 2. Pemrosesan Teks Data Komentar

Dataset yang telah terkumpul kemudian dilakukan pemrosesan teks dengan langkah *Case folding* atau penyamaan format teks, *Number removal* atau penghapusan nomor pada data teks, *whitespace removal* atau penghapusan karakter kosong, *punctuation removal* atau penghapusan tanda baca, *tokenization* atau pemecahan kalimat, *stop removal* dan *steming* atau penghilangan imbuhan yang dilakukan secara berturut-turut.

Tabel 1. Tabel Pemrosesan Teks

Komentar sebelum pemrosesan teks	Komentar setelah pemrosesan teks
@ummuakbarazka Halo. Produk jamu serbuk Gemuk Sehat ini tidak kami sarankan untuk dikonsumsi oleh ibu menyusui. Salam. ada efek samping ga?	mention halo produk jamu serbuk gemuk sehat ini tidak kami saran untuk konsumsi oleh ibu susu salam ada efek samping tidak

1.9 Pembobotan Kata

Pembobotan kata dilakukan menggunakan metode *Term Frequency – Inverse Document Frequency* (TF-IDF). Tabel berikut ini menunjukkan contoh hasil pembobotan salah satu komentar dalam dataset menggunakan *tools* yang digunakan.

Tabel 2. Pembobotan Kata Komentar

Komentar	Bobot Kata
mention waaah keren selalu sedia 	keren = 5.94542, mention = 0.278994, sedia = 3.86598, sidomunculstore = 1.85108, waaah = 5.25227,  = 5.94542

1.10 Klasifikasi *Naive Bayes*

Pengujian model *Naive Bayes* dalam pengklasifikasian dataset dilakukan dengan metode Cross Validation dengan *number folds* sebanyak 10. Dataset yang diuji adalah data training yang diambil dari 20% dari keseluruhan dataset yang dikumpulkan yaitu sebanyak 1906 komentar. *Score* evaluasi dari model *Naive Bayes* yang didapatkan adalah sebagai berikut:

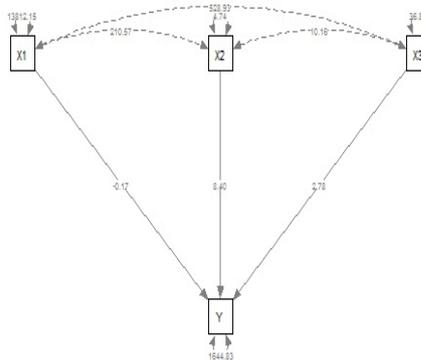
Tabel 3. *Testing* dan *Score* Data Training klasifikasi *Naive Bayes*

Model	AUC	CA	F1	Prec	Recall	MCC
<i>Naive Bayes</i>	0.867	0.559	0.684	0.961	0.559	0.368

Pada tabel 3 mengenai hasil evaluasi pada klasifikasi *Naive Bayes* menunjukkan nilai akurasi (CA) sebesar 0,599, nilai F1 sebesar 0,684, nilai *precision* sebesar 0,961, nilai *recall* 0,559, dan nilai MCC sebesar 0,368.

Istilah *world cloud* seperti yang ditunjukkan pada Gambar 5 merupakan tampilan visual kata yang sering banyak muncul dari 1906 dataset. Pada *world cloud* tersebut terdapat 10 kata yang paling sering muncul yaitu mention, sidomuncul, C, jalan, lebaranbarengsido, sidomunculstore, tolak, produk, yuk, dan linu.

1.11 Analisis Jalur *Trends Google* dan komentar produk herbal



Gambar 6. Diagram Plot *Google Trends* dan Data Komentar

Analisis jalur atau *path analysis* dilakukan menggunakan data kasual dari komentar positif, komentar netral, dan komentar negatif dipadukan dengan data *Google Trends* mengenai produk herba dan hasil analisis ditunjukkan seperti pada Gambar 6. Model yang dibuat bertujuan untuk melihat bagaimana relasi hubungan antara Y yaitu produk herba dengan X1 yaitu komentar positif, X2 yaitu komentar negatif, dan X3 yaitu komentar netral. Hasil menunjukkan bahwa pengaruh terbesar terdapat pada relasi X2 pada Y yakni dengan nilai sebesar 8.40 diikuti oleh X3 pada Y dengan nilai 2.78, dan terakhir X1 pada Y dengan nilai -0.17 .

KESIMPULAN

Dari keseluruhan data sejumlah 1906 komentar yang telah dikumpulkan melalui media media sosial *Instagram* data diklasifikasikan ke dalam 3 kategori yaitu komentar positif, netral, dan negatif. Model *Naive Bayes* yang diimplementasikan pada data training menghasilkan nilai *False Negatif* sebanyak 165 komentar. Dengan menggunakan nilai akurasi pada model sebesar 0,559, dataset diimplementasikan pada model *Naive Bayes* yang dibuat dan mengklasifikasikan data komentar sejumlah 875 sebagai komentar negatif, 147 data diklasifikasikan sebagai komentar netral, dan 884 data sebagai komentar positif. Kata mention yang paling banyak dilihat dari *world cloud* menunjukkan bahwa komentar instagram lebih banyak melakukan tak pada akun lain diikuti dengan kata sidomuncul. *Path Analysis* yang telah dilakukan menunjukkan bahwa komentar negatif yang paling banyak berpengaruh pada *trends google* mengenai produk herba diikuti oleh komentar netral dan terakhir komentar positif.

DAFTAR PUSTAKA

- [1] S. and V. K. and K. K. Bansal Malti and Verma, "Opinion Mining from Student Feedback Data Using Supervised Learning Algorithms," in *Third International Conference on Image Processing and Capsule Networks*, J. M. R. S. and S. F. Chen Joy Iong-Zong and Tavares, Ed., Cham: Springer International Publishing, 2022, pp. 411–418.
- [2] D. Suhartono, K. Purwandari, N. H. Jeremy, S. Philip, P. Arisaputra, and I. H. Parmonangan, "Deep neural networks and weighted word embeddings for sentiment analysis of drug product reviews," *Procedia Comput Sci*, vol. 216, pp. 664–671, 2023, doi: 10.1016/j.procs.2022.12.182.
- [3] T. Shaik, X. Tao, C. Dann, H. Xie, Y. Li, and L. Galligan, "Sentiment analysis and opinion mining on educational data: A survey," *Natural Language Processing Journal*, vol. 2, p. 100003, Mar. 2023, doi: 10.1016/j.nlp.2022.100003.
- [4] A. Hogenboom, A. Brojba-Micu, and F. Frascar, "The impact of word sense disambiguation on stock price prediction," *Expert Syst Appl*, vol. 184, Dec. 2021, doi: 10.1016/j.eswa.2021.115568.
- [5] R. Syukur Siregar, R. Ampuh Hadiguna, I. Kamil, and N. Nazir, "Permintaan Dan Penawaran Tanaman Obat Tradisional Di Provinsi Sumatera UTARA Demand and Supply Analysis of Traditional Medicinal Plants in Sumatera Utara," vol. 13, no. 1, pp. 50–59, 2020.
- [6] U. Nisa *et al.*, "Kualitas Hidup Pasien Batu Saluran Kemih Yang Menggunakan Ramuan Jamu Di Klinik Jejaring Saintifikasi Jamu," *Jurnal Tumbuhan Obat Indonesia*, vol. 14, no. 1, pp. 95–106, Jul. 2021, doi: 10.22435/jtoi.v14i1.4365.
- [7] H. Hartono and L. A. Kusumastuti, "Tingkat Kepatuhan Penggunaan Ramuan Jamu Saintifik Hiperglikemia pada Pasien Diabetes Melitus di Rumah Riset Jamu Hortus Medicus B2P2TOOT Tawangmangu," *Jurnal Farmasi (Journal of Pharmacy)*, vol. 8, no. 1, Oktober, pp. 29–37, Nov. 2019, doi: 10.37013/jf.v1i8.77.
- [8] M. A. Marwati Amidi, "Pengaruh Budaya, Persepsi, Dan Kepercayaan Terhadap Keputusan Pembelian Obat Herbal," *Jurnal Ilmu Manajemen*, no. Vol 7, No 2 (2018): Jurnal Ilmu Manajemen, pp. 168–180, 2018, [Online]. Available: https://jurnal.um-palembang.ac.id/ilmu_manajemen/article/view/1567/1315
- [9] R. Eri, N. Kurniawati, J. Sukoco, A. Windriya, and D. Iskandar, "Pelatihan Pemasaran Produk Homemade Melalui Sosial Media," *JPPM (Jurnal Pengabdian dan Pemberdayaan Masyarakat)*, vol. 4, p. 1, Mar. 2020, doi: 10.30595/jppm.v4i1.4655.
- [10] S. H. U. V. A. Q. V. P. L. Seprillina, "Peningkatan Omset Penjualan Melalui Media Sosial Pada Usaha Kecil Menengah Alas Kaki Berbahan Kulit," *Jurnal KARINOV*, no. Vol 2, No 1 (2019): Januari, pp. 57–61, 2019, [Online]. Available: <http://journal2.um.ac.id/index.php/jki/article/view/8267/3932>
- [11] Z. A. A. T. Z. A. W. N. E. N. N. A. F. D. S. I. Cahyaningrum, "Pemanfaatan Media Sosial dalam Pemasaran Produk UMKM di Kelurahan Sidokumpul, Kabupaten Gresik," *Jurnal Ilmu Komunikasi*, no. Vol. 10 No. 1 (2020): April, pp. 17–31, 2020, [Online]. Available: <http://jurnalfdk.uinsby.ac.id/index.php/JIK/article/view/359/320>
- [12] I. A. M. Y. Toni, "Analisis Sentimen Terhadap Tempat Kuliner Ayam Gedebuk Dari Komentar Pengunjung Dengan Menggunakan Metode Naïve Bayes Classifier," *JBASE - Journal of Business and Audit Information Systems*, no. Vol 3, No 1 (2020): JBASE-Journal of Business and Audit Information Systems, 2020, [Online]. Available: <https://journal.ubm.ac.id/index.php/jbase/article/view/2062/1687>
- [13] W. Parasati, F. Abdurrachman Bachtiar, and N. Y. Setiawan, "Analisis Sentimen Berbasis Aspek pada Ulasan Pelanggan Restoran Bakso President Malang dengan Metode Naïve Bayes Classifier," 2020. [Online]. Available: <http://j-ptiik.ub.ac.id>

- [14] M. A. Rofiqi, Abd. C. Fauzan, A. P. Agustin, and A. A. Saputra, "Implementasi Term-Frequency Inverse Document Frequency (TF-IDF) Untuk Mencari Relevansi Dokumen Berdasarkan Query," *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 1, no. 2, pp. 58–64, Dec. 2019, doi: 10.28926/ilkomnika.v1i2.18.
- [15] E. Dwianto and M. Sadikin, "Analisis Sentimen Transportasi Online pada Twitter Menggunakan Metode Klasifikasi Naïve Bayes dan Support Vector Machine," *Format : Jurnal Ilmiah Teknik Informatika*, vol. 10, p. 94, Feb. 2021, doi: 10.22441/format.2021.v10.i1.009.
- [16] M. W. A. Putra, Susanti, Erlin, and Herwin, "Analisis Sentimen Dompot Elektronik Pada Twitter Menggunakan Metode Naïve Bayes Classifier," *IT Journal Research and Development*, vol. 5, no. 1, pp. 72–86, Jul. 2020, doi: 10.25299/itjrd.2020.vol5(1).5159.
- [17] F. Rustam, A. Mehmood, M. Ahmad, S. Ullah, D. M. Khan, and G. S. Choi, "Classification of Shopify App User Reviews Using Novel Multi Text Features," *IEEE Access*, vol. 8. Institute of Electrical and Electronics Engineers Inc., pp. 30234–30244, 2020. doi: 10.1109/ACCESS.2020.2972632.
- [18] M. H. S. S. A. Hoyyi, "Analisis Jalur Terhadap Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif (Ipk) Mahasiswa Statistika Undip," *Jurnal Gaussian*, no. Vol 4, No 1 (2015): *Jurnal Gaussian*, pp. 61–70, 2015, [Online]. Available: <https://ejournal3.undip.ac.id/index.php/gaussian/article/view/8146/7922>