

KLASIFIKASI TRENDING TOPIK DI MEDIA SOSIAL X MENGGUNAKAN ALGORITMA NAÏVE BAYES BERBASIS PARTICLE SWARM OPTIMIZATION

Endang Sri Palupi

Teknologi Informasi, Universitas Bina Sarana Informatika
Jalan Kramat Raya No. 98 Jakarta Pusat
endang.epl@bsi.ac.id

ABSTRAK

Media sosial X salah satu platform media sosial yang banyak digunakan untuk memposting berbagai macam kejadian dan peristiwa viral. Postingan berisi isu hangat, berita terbaru, atau peristiwa viral yang menjadi perhatian baik di dalam negeri maupun luar negeri. Trending topik di media sosial X adalah daftar kata kunci, frasa, atau tagar, viral, dan banyak diperbincangkan dalam kurun waktu tertentu. Oleh karena itu penulis ingin mengklasifikasi trending topik sehingga bisa mendapatkan informasi yang lebih baik, dengan membagi beberapa trending topik sesuai pembahasannya. Penelitian ini datanya diambil dari media sosial X pada Januari 2024 sampai dengan April 2025 dengan teknik crawling yaitu proses otomatis untuk mengumpulkan data postingan, informasi user, dan interaksi. Kemudian metode text mining digunakan untuk mengekstraksi informasi berharga dari unggahan yang tidak terstruktur. Metode yang digunakan pada penelitian ini menggunakan pemodelan algoritma naïve bayes berbasis particle swarm optimization dan dilakukan validasi data menggunakan k-fold cross validation dengan uji nilai fold = 10. Dari pengujian yang telah dilakukan menggunakan framework rapidminer mendapatkan hasil akurasi sebesar 97% dengan class precision pred NO : 99,58% dan pred YES 93,12%, class recall true NO 95,60% dan true YES 99,33%. Dari hasil penelitian ini maka klasifikasi menggunakan metode algoritma naïve bayes berbasis PSO merupakan pengolahan data klasifikasi yang baik.

Kata kunci : klasifikasi, naïve bayes, X

1. PENDAHULUAN

Topik yang sedang tren di X selalu berubah, tetapi beberapa contoh tren di masa lalu antara lain #IndonesiaGelap, "Peringatan Darurat" terkait RUU Pilkada, dan topik terkait artis atau acara tertentu, seperti Aura Farming dan Rayyan Arkan Dikha. Anda dapat melihat topik yang sedang tren secara langsung dengan membuka bagian "Tren" di tab "Jelajahi" pada aplikasi seluler, atau di halaman beranda/pencarian pada versi desktop, seperti yang dijelaskan di Pusat Bantuan X. Topik yang Sedang Tren adalah fitur layanan di X yang menyediakan informasi tentang topik terhangat atau informasi terbaru yang sedang dibahas.

Media sosial X adalah salah satu platform media sosial untuk netizen berinteraksi dengan berbagi pesan singkat, mengikuti akun, melalui balasan, retweet, dan menyukai, serta menjadi tempat untuk percakapan real-time dan berita yang sedang trending saat ini. Dalam berkomunikasi sehari-hari, pengguna media sosial cenderung menggunakan bahasa yang lebih santai untuk menciptakan keakraban antar pengguna sehingga interaksi menjadi lebih menarik. Bahasa gaul digunakan sebagai ekspresi rasa keakraban para netizen di dunia maya.[1]

Klasifikasi merupakan teknik pengolahan data dengan cara mengelompokkan atau penggolongan data secara sistematis ke dalam kelas – kelas sesuai dengan karakteristik dan kesamaan ciri yang dimiliki oleh masing-masing data. Dalam pengolahan datanya, data yang akan diolah tidak memiliki persyaratan tertentu.

Singkatnya semua data dapat diklasifikasi. Teknik klasifikasi data tersebut dapat menggunakan berbagai macam algoritma dalam data mining. Terdapat banyak algoritma data mining yang dapat digunakan dalam klasifikasi, seperti Decision Tree, Support Vector Machine, Neural Network, dan K-Nearest Neighbor serta Naive Bayes yang seperti penulis gunakan dalam penelitian ini.[2] Klasifikasi adalah tipe analisis data yang dapat membantu menentukan kelas label dari sampel yang ingin di klasifikasi. Klasifikasi merupakan Metode supervised learning, metode yang mencoba menemukan hubungan antara atribut masukan dan atribut target. Tujuan klasifikasi untuk meningkatkan hasil akurasi yang diperoleh dari pengolahan data.[3] Penerapan analisis data mining, khususnya teknik klasifikasi banyak digunakan karena kesederhanaannya. Teknik klasifikasi melibatkan dua tahap utama pertama, mengembangkan model menggunakan data pelatihan kedua, mengevaluasi model tersebut dengan data pelatihan yang sama. Ada banyak metode dalam teknik klasifikasi, termasuk algoritma statistik, analisis korelasi, analisis regresi, Naïve Bayes, algoritma berbasis jarak, pendekatan sederhana, k-nearest neighbors, pohon keputusan, jaringan syaraf tiruan, dan algoritma berbasis aturan. [4]

Penelitian ini bertujuan untuk menguji algoritma naïve bayes untuk mengklasifikasikan trending topik yang ada pada media sosial X, sehingga dapat mengetahui tingkat akurasi dalam mengklasifikasikan topik yang sedang trending pada media sosial X.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Di jaman digitalisasi saat ini, kecepatan arus informasi semakin meningkat, dengan meluasnya penggunaan internet di masyarakat dunia. Media sosial "X" menjadi indikator yang jelas tentang bagaimana masyarakat mengandalkan platform ini untuk mencari informasi yang akurat dan terpercaya, termasuk di kalangan mahasiswa program studi Ilmu Perpustakaan dan Informasi, Universitas Negeri Padang. Meskipun awalnya media sosial seperti "X" lebih banyak digunakan untuk interaksi sosial dan hiburan, kini penggunaannya telah meluas sebagai sumber informasi yang cukup penting. Dalam penelitian ini, Qalbina dan rekan-rekannya menggunakan metode penelitian kuantitatif dengan pendekatan deskriptif. Objek penelitian terdiri dari 80 mahasiswa program studi Ilmu Perpustakaan dan Informasi, dengan metode pengambilan data penyebaran kuesioner kepada responden sesuai kriteria penelitian. Hasil penelitian ini berupa skala interval 3,41-4,20, di mana penggunaan fitur topik yang sedang tren mampu memenuhi kebutuhan informasi bagi mahasiswa dengan baik. Perbedaan dengan penelitian ini, metode pengambilan datanya langsung dari media sosial X dengan teknik crawling dan text mining. Dalam pengolahan datanya penulis menggunakan data mining untuk klasifikasi algoritma naïve bayes berbasis PSO dengan menggunakan framework Rapidminer [5]

Dalam penelitian Wijayanti dan rekan-rekannya, salah satu metode yang dapat mengakomodasi hal ini adalah metode klasifikasi sentimen berbasis Support Vector Machine (SVM) yang dikombinasikan dengan Term Frequency-Inverse Document Frequency (TF-IDF). Menganalisis pengaruh topik yang sedang tren di media sosial, khususnya TikTok, terhadap kepercayaan konsumen dan perilaku belanja daring sangatlah penting karena dapat memberikan wawasan tentang sikap konsumen. Analisis klasifikasi sentimen menggunakan regresi linier menunjukkan bahwa topik yang sedang trending memiliki pengaruh yang signifikan terhadap kepercayaan konsumen, yang selanjutnya mempengaruhi niat beli dan perilaku belanja digital. Oleh karena itu, era digital dengan pesatnya pertumbuhan media sosial menunjukkan bahwa sentimen viral sangat penting dalam keputusan pembelian konsumen. Sedangkan penulis dalam penelitian ini khusus hanya trending topik dalam media sosial X saja, yang berisi berita dan sesuatu hal yang sedang viral. Media sosial X berbeda dengan tiktok yang sudah terafiliasi dengan tokopedia, sehingga banyak netizen yang ingin berbelanja di media sosial tersebut. [6]

Pada tahun 2023, Catur Arpal dan rekan-rekannya melakukan penelitian mengenai keterlibatan buzzer dalam penyebaran berita dan konten hoax. Penelitian ini mengkaji pola algoritma akun-akun yang terlibat dalam penyebaran berita tersebut. Penelitian ini mengembangkan klasifikasi buzzer berdasarkan

pola dan karakteristik yang diperoleh dari akun-akun tersebut di media sosial X. Varian Gaussian dari algoritma naïve bayes digunakan untuk klasifikasi akun buzzer dan non-buzzer. Hasil penelitian ini menunjukkan hasil model klasifikasi akun buzzer dengan nilai akurasi 80%.Perbedaannya dengan penelitian ini, penulis mengklasifikasi trending topik dalam kurun waktu tertentu di media sosial X dan menggunakan optimisasi PSO untuk meningkatkan nilai akurasi algoritma naïve bayes.[7]

Pada penelitian selanjutnya yang berjudul Analisis Sentimen Maskapai Penerbangan AS pada Media Sosial Twitter/X membandingkan algoritma data mining, penelitian ini menggunakan dataset dari media sosial Twitter Maskapai Penerbangan AS dengan menggunakan dua vektorisasi yaitu Bag Of Words dan Tf-Idf sebagai pembandingan untuk mendapatkan nilai akurasi. Untuk evaluasi data tweet menggunakan metode algoritma data mining seperti Multi Nominal Naïve Bayes, ensemble voting untuk menggambarkan konsistensi akurasi. Berdasarkan hasil perhitungan masing-masing algoritma data mining diketahui bahwa nilai akurasi dataset Twitter Maskapai Penerbangan AS pada analisis sentimen cukup tinggi yaitu sebesar 0,95 atau 95% dengan menggunakan metode data mining algoritma SVM. Sedangkan vektorisasi dataset Twitter Maskapai Penerbangan AS dengan bag of words mempunyai nilai akurasi tertinggi sebesar 0,92 atau 92% dengan menggunakan metode data mining algoritma Multinomial Naïve Bayes. Maka dari hasil penelitian tersebut hasil akurasi lebih baik menggunakan metode data mining algoritma SVM dibandingkan algoritma naïve bayes.Sedangkan penulis dalam penelitian ini mengklasifikasi hanya menggunakan satu algoritma yaitu naïve bayes dan menggunakan optimisasi PSO untuk mendapatkan nilai akurasi yang lebih tinggi.[8]

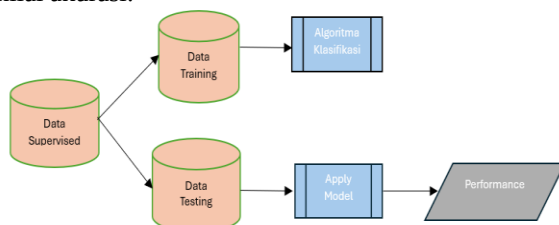
Penelitian yang ditulis oleh Aulia Cisatra dan rekan-rekannya, berjudul "Mengidentifikasi Topik Hangat di Media Berita Menggunakan Alokasi Dirichlet Laten", menggunakan pendekatan pemrosesan bahasa alami. Model alokasi dirichlet laten akan digunakan untuk analisis dan klasifikasi berita berdasarkan topik. Data teks digital yang digunakan berasal dari berbagai portal berita nasional melalui akses antarmuka pemrograman aplikasi (API). Melalui penambangan teks dan pembelajaran mesin, ekstraksi otomatis topik utama dari data teks besar yang tidak terstruktur dapat diidentifikasi dengan cepat dan akurat. Hasil penelitian menunjukkan bahwa tiga topik terpopuler berdasarkan dominasi token, secara berurutan, adalah Komisi Pemberantasan Korupsi (KPK) (36,6%), calon wakil presiden (32,8%), dan Jakarta (30,6%). Sedangkan penelitian ini mengklasifikasi trending topik khusus di media sosial X dan menggunakan metode Naïve Bayes untuk mengklasifikasi sebagai salah satu algoritma dalam data mining. [9]

2.2. Data Mining

Data mining adalah proses menggunakan teknik statistik, matematika, dan kecerdasan buatan untuk menemukan pola, informasi dan tren yang berharga yang tersembunyi di dalam data yang besar. Tujuan utama dari data mining adalah mengungkap informasi yang tidak terlihat secara langsung atau dapat diidentifikasi dengan mudah melalui analisis rutin sehingga dapat membantu pengambilan keputusan diberbagai bidang. [10] *Data mining* adalah proses penyaringan data yang memanfaatkan kumpulan data yang cukup besar melalui beberapa proses untuk mengekstrak informasi berharga dari data tersebut. Perusahaan besar dapat menggunakan *data mining* untuk mengolah serta memeriksa data untuk memperoleh informasi yang dapat mendukung pengambilan keputusan dan meningkatkan proses bisnis sehingga tercapai tujuan perusahaan.[11]

2.3. Naïve Bayes

Naïve Bayes Classifier adalah metode klasifikasi yang berakar pada teorema Bayes merupakan salah satu metode dalam data mining. Naive bayes merupakan metode pengklasifikasian berdasarkan probabilitas posterior menggunakan probabilitas prior dan probabilitas likelihood dengan asumsi setiap fitur berkontribusi secara independen terhadap kelas, walaupun kenyataannya tidak selalu begitu. Keunggulan dari metode ini cepat dan efisien, mudah di implemetasikan, bisa bekerja dengan data kuantitatif ataupun kualitatif. Sedangkan kekurangannya adalah asumsi independen fitur seringkali tidak akurat dalam dunia nyata. Tujuan dari metode Naïve Bayes adalah untuk menemukan probabilitas ketika sudah mengetahui probabilitas tertentu lainnya. Hasil dari perhitungan data mining menggunakan metode klasifikasi Naïve Bayes akan makin berguna jika penyajiannya baik dan bisa dipahami dengan baik oleh penerima data.[10] Pada gambar 1 proses klasifikasi dimana dataset displit menjadi data training dan data testing kemudian pemodelan menggunakan framework Rapidminer dengan menggunakan algoritma Naïve bayes dengan apply model untuk mendapatkan performance yaitu nilai akurasi.



Gambar 1. Pemodelan Klasifikasi Data Mining

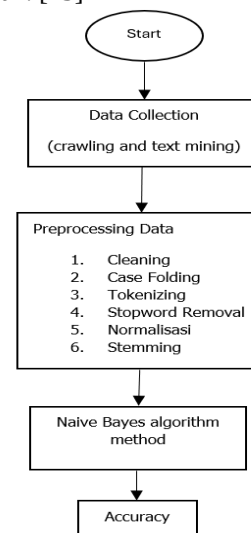
Klasifikasi data mining adalah proses mengkategorikan data ke dalam kelas atau label yang telah ditentukan sebelumnya untuk membangun model prediktif. Ini merupakan bagian dari supervised learning, di mana model dilatih menggunakan data yang sudah memiliki label untuk memprediksi label

data baru yang belum diketahui. Tujuannya adalah mengorganisasi data agar lebih mudah dipahami dan digunakan untuk membuat keputusan atau prediksi di masa mendatang. Klasifikasi ialah sebuah prosedur mendeteksi kumpulan pola maupun guna yang menggambarkan dan memisahkan kasta data yang satu dengan yang yang lain untuk melaporkan entitas itu masuk pada golongan terpilih yang telah ditetapkan.[12]

3. METODE PENELITIAN

3.1. Jenis Penelitian

Jenis penelitian ini bersifat kuantitatif, mengambil data trending topik media sosial X dalam kurun waktu Januari 2024 hingga April 2025 menggunakan teknik crawling dan text mining. Penelitian kuantitatif merupakan pendekatan penelitian yang beraliran positivisme. Pendekatan kuantitatif menggunakan alat uji statistik dan matematika yang bisa dihitung dengan angka. Sementara itu, pendekatan kualitatif lebih didasarkan pada penalaran yang logis dan pemahaman interpretasi objek penelitian. [13]



Gambar 2. Tahapan Penelitian

3.2. Pengumpulan Data

Teknik pengikisan data dan penambangan teks digunakan untuk mengumpulkan data dari platform media sosial X, topik-topik yang sedang tren antara Januari 2024 dan April 2025. Penyesuaian parameter untuk pengikisan data dari X merupakan langkah krusial dalam mengekstrak data berdasarkan kata kunci, tagar, tanggal dan waktu, lokasi, URL, dll. Setelah menjalankan alat pengikisan, data dapat diekspor ke berbagai format berkas (xlsx, CSV, JSON, dll.). Langkah terakhir adalah menganalisis data yang diekspor untuk mendapatkan wawasan tentang topik-topik yang sedang tren. [14]

3.3. Preprocessing Data

Preprocessing data merupakan proses cukup penting dalam analisis data mining. Proses ini bertujuan untuk membersihkan data, memformat

ulang, dan mempersiapkan data agar analisisnya lebih mudah dan akurat. Setelah preprocessing data selesai, data siap digunakan dalam analisis data mining dengan pemodelan menggunakan framework rapidminer. Data ini dapat diterapkan pada teknik data mining klasifikasi untuk mendapatkan wawasan baru yang dapat membantu pengambilan keputusan.[15] Dalam studi ini, penulis melakukan preprocessing data sebagai berikut:

- Case Folding: proses merubah semua huruf kapital dalam dataset menjadi huruf kecil.
- Noise Removal: proses menghilangkan noise dari dataset, seperti hastag, angka, tanda baca, URL, dan atribut lain yang mengandung nilai yang hilang.
- Tokenizing: proses memecah urutan kalimat menjadi beberapa kata yang disebut token.
- Stopword Removal: proses menghilangkan kata-kata yang tidak berguna dari dataset menggunakan stopwords.
- Stemming: proses mengubah setiap kata dalam dataset ke bentuk dasarnya.

3.4. Metode Algoritma Naïve Bayes

Pengklasifikasi Naïve Bayes adalah metode klasifikasi yang memiliki beberapa tahapan: data training dan data testing (splitting), algoritma, model, data uji, dan pengujian, yang semuanya pada akhirnya menghasilkan akurasi yang akurat. Pemodelan algoritma Naïve Bayes mudah dan cepat dibangun dan sangat berguna untuk data yang sangat besar. (Martantoh & Yanih, 2022) dalam penelitian penulis melakukan pemodelan menggunakan Rapidminer studio versi 10.1

3.5. Akurasi

Akurasi adalah proporsi semua klasifikasi yang benar, baik positif maupun negatif. Akurasi didefinisikan secara matematis sebagai berikut: $Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$.

4. HASIL DAN PEMBAHASAN

Hasil pengumpulan data trending topik media sosial X dari Januari 2024 hingga April 2025 adalah sebagai berikut. Berdasarkan hasil tahap preprocessing data sebelumnya, data yang diperoleh dapat digunakan untuk menyusun bag of words. Hasilnya ditunjukkan pada Tabel 1.

Tabel 1. Bag Of Words

No	TOPIK	TWEET	TOKEN	TERM UNIK
1	Politik	44,67	39,81	40,72
2	Issue terkini	50,44	51,78	49,78
3	Ekonomi	40,53	39,21	38,11
4	Pendidikan	37,45	40,31	35,22
5	Olahraga	42,89	40,28	39,72
6	Selebriti	49,34	45,25	44,52
7	Kesehatan	37,51	37,91	35,21
8	Hiburan	42,78	41,74	40,32

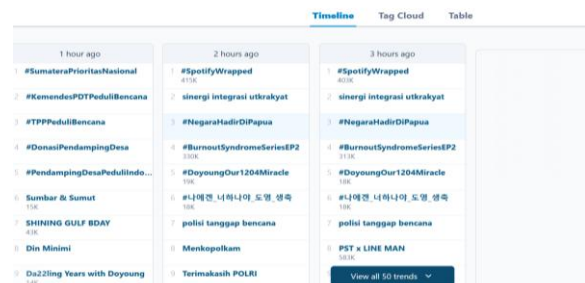
Tabel 2 menunjukkan hasil pelabelan data X berdasarkan kategori.

Tabel 2. Hasil Labeling

Label	keyword
Politik	Jokowi, Capres, Pemilu, Menteri
Issue terkini	Arhan, Sin tae yong, #kaburajadulu
Ekonomi	#Bitcoin, inflasi, rupiah, dollar
Pendidikan	Guru, sekolah, #SMBPTN2024, bully
Olahraga	Bola, kalah telak, piala dunia, PSSI
Selebriti	Arhan, zizie, mamagigi, #raffiahmad
Kesehatan	Jantung, penyakit, covid19
Hiburan	film, musik, #KPOP, #globalTV
Teknologi	Microsoft, google, 5G, gadget
Budaya	Sejarah, festival, baju daerah
Motivasi	Quot, motivasi, motivator, kutipan

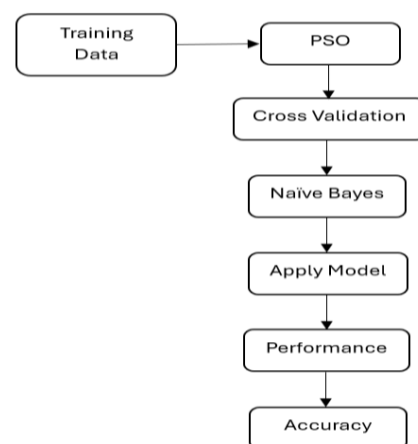
Topik yang sedang tren di media sosial X juga bisa dicek di <https://trends24.in/indonesia/> Begini tampilannya :

Indonesia X (Twitter) Trends for last 24 hours



Gambar 3. Trending Topik X Indonesia

4.1. Pemodelan Algoritma Naïve Bayes Berbasis Particle Swarm Optimization

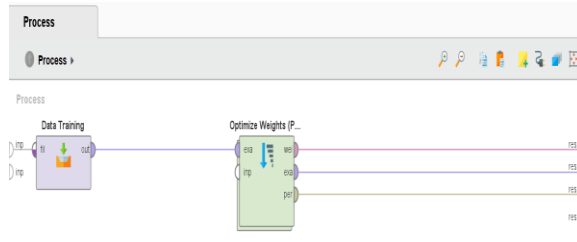


Gambar 4. Pemodelan Algoritma Naïve Bayes berbasis Particle Swarm Optimization

Gambar 4 adalah tahapan pemodelan menggunakan algoritma Naïve Bayes berbasis Particle Swarm Optimization (PSO) menggunakan RapidMiner.

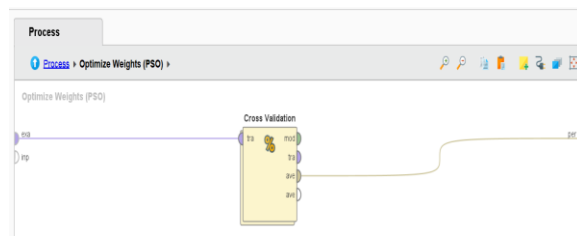
- Data training diproses menggunakan PSO untuk mengoptimalkan akurasi. Penelitian sebelumnya telah menunjukkan bahwa PSO, sebuah algoritma optimasi berbasis populasi yang terinspirasi oleh

perilaku sosial swarm, memungkinkan eksplorasi ruang solusi yang lebih efisien dan efektif dibandingkan dengan metode tradisional. Hasilnya menunjukkan bahwa penerapan PSO pada ANN secara signifikan meningkatkan metrik kinerja model, termasuk akurasi, presisi, recall, dan skor F1, dibandingkan dengan metode dasar. [16] berikut pemodelan data training menggunakan PSO terlihat pada gambar 5.



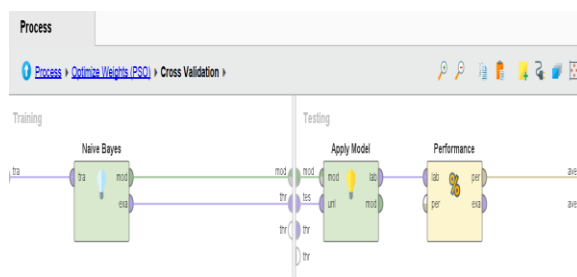
Gambar 5. Pemodelan Data Training

- b. Langkah selanjutnya melibatkan *cross validation*, sebuah teknik evaluasi yang digunakan dalam pembelajaran machine learning untuk menilai seberapa baik model akan bekerja pada data baru dengan splitting data sehingga memberikan kinerja yang lebih akurat dan mencegah *overfitting* (model terlalu hafal data latih). *Cross Validation* membantu memperkirakan semua data yang mungkin dengan membuat beberapa model, sehingga hasilnya lebih andal. [17] Penggunaan *cross validation* terlihat pada gambar 6.



Gambar 6. Cross Validation

- c. Pemodelan menggunakan algoritma Naïve Bayes, kemudian apply model dan performa untuk mendapatkan hasil akurasi. Tujuan apply model adalah untuk menghubungkan data testing dan memeriksa akurasi serta kesalahan model. Performance (klasifikasi) digunakan untuk memeriksa tingkat akurasi model. [18] Pemodelan naïve bayes terlihat pada gambar 7.



Gambar 7. Pemodelan Algoritma Naïve Bayes

Tabel 3. Hasil akurasi Algoritma Naïve Bayes Berbasis PSO

Accuracy : 97%	True NO	True YES	class precision
Pred NO	239	1	99.58%
Pred YES	11	149	93.12%
class recall	95.60%	99.33%	

Dari tabel 3 berikut perhitungan manual akurasi algoritma Naïve Bayes berbasis PSO :

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

$$Accuracy = \frac{239 + 149}{239 + 1 + 11 + 149} \times 100\%$$

$$Accuracy = \frac{388}{400} \times 100\% = 0,97 = 97\%$$

5. KESIMPULAN DAN SARAN

Kesimpulan penelitian ini bertujuan untuk mengklasifikasikan topik yang sedang tren di media sosial X pada bulan Januari 2024 – April 2025. Trending topik adalah topik yang sedang populer dan banyak dibicarakan di media sosial dalam kurun waktu tertentu, seperti berita terkini, acara TV, isu sosial, atau konten viral. Topik ini menjadi tren karena mendapat atensi tinggi dari pengguna, seringkali ditandai dengan tagar. Metode penelitian ini menggunakan algoritma naïve bayes berbasis particle swarm optimization, dengan hasil akurasi sebesar 97% dengan kelas presisi pred NO: 99,58% dan pred YES 93,12%, kelas recall true NO 95,60% dan true YES 99,33%. Dengan hasil penelitian ini, klasifikasi menggunakan metode algoritma naïve bayes berbasis PSO merupakan pengolahan data klasifikasi yang baik. Untuk penelitian selanjutnya, dapat digunakan lebih dari satu algoritma untuk perbandingan, dan dapat digunakan lebih banyak atribut dengan berbagai topik di media sosial X maupun media sosial lainnya.

DAFTAR PUSTAKA

- [1] M. A. Faidh *et al.*, “Peran Media Sosial X Dalam Perkembangan Komunikasi Di Era Digital,” *Konsensus J. Ilmu Pertahanan, Huk. dan Ilmu Komun.*, vol. 1, no. e-ISSN : 3048-1325, pp. 43–51, 2024, doi: 10.62383/konsensus.v1i6.433.
- [2] N. A. Ramadhani and H. A. Rosyid, “Algoritma – Algoritma Data Mining untuk Klasifikasi Data,” *J. Inov. Tek. dan Edukasi Teknol.*, vol. 2, no. 12, pp. 550–556, 2022, doi: 10.17977/um068v2i122022pxxx-xxx.
- [3] B. I. Nugroho and Z. Arif, “Tinjauan Pustaka Sistematis : Penerapan Data Mining Metode Klasifikasi Untuk Menganalisa Penyalahgunaan Sosial Media,” *J. Sist. Inf. dan Teknol. Perad.*, vol. 3, no. 2, pp. 46–51, 2022.
- [4] A. Naufal, W. Zain, and A. Tholib, “Klasifikasi Data Mining di Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Sistem Informasi Fakultas

- Teknik Universitas Nurul Jadid,” *SKANIKA Sist. Komput. dan Tek. Inform.*, vol. 7, no. E-ISSN: 2721-4788, pp. 204–213, 2024.
- [5] G. S. Qalbina and R. Asmara, “Pemanfaatan Fitur Trending Topic Twitter ‘X’ Sebagai Upaya Pemenuhan Kebutuhan Informasi Mahasiswa Program Studi Perpustakaan Dan Ilmu Informasi Universitas Negeri Padang,” *J. Inov. dan Pengabd. Kpd. Masy.*, vol. 3, no. p-ISSN : 2964-4992, pp. 816–824, 2024, doi: <https://doi.org/10.58578/alldyas.v3i2.3140>.
- [6] L. Wijayanti, N. E. Febriyanti, and I. Suarti, “Analisa Pengaruh Trending Topik Media Sosial terhadap Kepercayaan dan Perilaku Belanja Online Konsumen Menggunakan Metode Sentiment Classification Berbasis SVM,” *JUTEKNIN Telecommun. Electron. Autom. Robot. Mechatronics, Inf. Syst. Comput. Network, Comput. Electron. Control Power*, vol. 1, no. 1, pp. 1–9, 2025.
- [7] C. A. Perkasa, A. A. Arifiyanti, and A. Salim, “Klasifikasi Akun Buzzer Pada Twitter Menggunakan Algoritma Naive Bayes,” *J. Ilm. Tek. Inform. dan Komun.*, vol. 3, no. 1, pp. 1–12, 2023.
- [8] A. Fauzi and A. H. Yunial, “Analisis Sentimen US Airline Pada Media Sosial Twitter / X Menggunakan Perbandingan Algoritma Data Mining,” *J. Edukasi dan Penelit. Inform.*, vol. 10, no. 2, pp. 277–286, 2024.
- [9] A. Cisatra, D. Daris, M. Ansori, D. Nasywa, F. Afiqah, and S. Ramadhani, “Identifikasi Topik Hangat di Media Berita Menggunakan Latent Dirichlet Allocation,” *J. Inf. Eng. Educ. Technol.*, vol. 08, no. ISSN : 2549-869X, pp. 14–17, 2024.
- [10] E. Mardiani, N. Rahmansyah, I. Kurniati, and N. Matondang, “Membandingkan Algoritma Data Mining Dengan Tools Orange untuk Social Economy,” *Digit. Transform. Technol.*, vol. 3, no. 2, pp. 686–693, 2023, doi: <https://doi.org/10.47709/digitech.v3i2.3256>.
- [11] M. F. Haryanti, A. Fauzi, A. A. Jelita, and A. Setiyowati, “Pengaruh Data Mining , Strategi Perusahaan Terhadap Laporan Kinerja Perusahaan,” *J. Manajemen dan Bisnis*, vol. 3, no. 1, pp. 71–90, 2024.
- [12] A. A. Wijaya, Kurniadi, and Hendri, “Klasifikasi Data Mining Dalam Menentukan Produk Vapor-Juice Terlaris Menggunakan Naive Bayes Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM),” *J. Inform. Dan Rekayasa Komput.*, vol. 4, no. September, pp. 1225–1235, 2024, doi: 10.33998/jakakom.v4i2.
- [13] S. Thaha and S. Jambi, “Pendekatan penelitian kuantitatif dan kualitatif serta tahapan penelitian,” *J. Genta Mulia*, vol. 15, no. 1, pp. 82–92, 2024, [Online]. Available: <https://ejournal.stkipbbm.ac.id/index.php/gm>
- [14] Farhad Pashaei, “Pengikisan Twitter (X) pada tahun 2025: Panduan Utama,” *Rapid SeedBox*, 2025. <https://www.rapidseedbox.com/id/blog/masterin-g-twitter-scraping>
- [15] A. Agung, A. Daniswara, and I. K. D. Nuryana, “Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru,” *J. Informatics Comput. Sci.*, vol. 05, pp. 97–100, 2023.
- [16] M. Jamhuri and T. Utomo, “Penggunaan Particle Swarm Optimization Pada Jaringan Syaraf Tiruan Untuk Klasifikasi Sinyal Radar,” *Math. Appl. J.*, pp. 192–201, 2024.
- [17] PT Revolusi Cita Edukasi, “Apa itu Cross Validation?,” 2025. <https://www.revou.co/kosakata/cross-validation>
- [18] A. Ifitah and R. Setyadi, “Penerapan Algoritma C . 45 Untuk Analisis Pengadaan Peralatan dan Mesin Kantor,” *J. Inf. Syst. Res.*, vol. 4, no. 2, pp. 434–442, 2023, doi: 10.47065/josh.v4i2.2673.