PEMETAAN TOPIK, TEKNIK, DAN ALGORITMA DATA MINING DALAM PENELITIAN KEIMIGRASIAN: TINJAUAN PUSTAKA SISTEMATIS

I Nyoman Bayu Satrya Mataram[1], Catur Susaningsih[2], Isidorus Anung Prabadhi[3]
[1,2,3]Manajemen Teknologi Keimigrasian, Politeknik Imigrasi dan Pemasyarakatan
bayuinyoman@gmail.com[1], catursusan7@gmail.com[2], anung53@gmail.com[3]

ABSTRACT

The growing digitalization of immigration services has increased the availability of administrative and public-interaction data, yet the distribution of research topics and the analytical methods used in Indonesian immigration studies has not been mapped systematically. This study maps topics, analytical techniques, and data-mining algorithms through a systematic literature review informed by PRISMA 2020. Searches in Google Scholar and OpenAlex identified 259 records; 29 candidate publications were reviewed, and 19 primary studies published from 2018 to 2026 were included. Data were extracted on research topic, data source, analytical task, algorithm, and principal result, then synthesized through frequency tabulation and narrative analysis. Public sentiment was the dominant topic with 13 studies, followed by Indonesian migrant workers with two studies. Passport applications, residence permits, visitor patterns, and security monitoring each appeared once. Classification was the leading analytical task with 13 studies, while Naive Bayes was the most frequently used algorithm with 11 applications. The evidence remains concentrated on publicly accessible reviews and social-media data, whereas operational immigration data and cross-location validation are limited. Future studies should diversify topics, establish privacy-governed access to anonymized operational data, compare stronger baselines, and report evaluation procedures more transparently.

Keywords: data mining, immigration, research mapping, PRISMA, systematic literature review, Naive Bayes

ABSTRAK

Digitalisasi layanan keimigrasian meningkatkan ketersediaan data administratif dan data interaksi publik, tetapi persebaran topik serta metode analisis dalam penelitian keimigrasian Indonesia belum dipetakan secara sistematis. Penelitian ini bertujuan memetakan topik, teknik analisis, dan algoritma *data mining* melalui tinjauan pustaka sistematis yang mengacu pada PRISMA 2020. Pencarian pada Google Scholar dan OpenAlex menghasilkan 259 rekaman; sebanyak 29 studi kandidat ditelaah dan 19 studi primer terbitan 2018-2026 dimasukkan dalam sintesis. Data diekstraksi berdasarkan topik, sumber data, tugas analitik, algoritma, dan hasil utama, kemudian dianalisis melalui tabulasi frekuensi dan sintesis naratif. Sentimen masyarakat menjadi topik dominan dengan 13 studi, diikuti Pekerja Migran Indonesia sebanyak dua studi. Paspor, izin tinggal, pola pengunjung, serta pengawasan keamanan masing-masing muncul satu kali.

Klasifikasi menjadi teknik utama dengan 13 studi, sedangkan *Naive Bayes* merupakan algoritma yang paling sering digunakan dengan 11 penerapan. Bukti masih terkonsentrasi pada ulasan publik dan media sosial, sedangkan penggunaan data operasional serta validasi lintas lokasi masih terbatas. Penelitian selanjutnya perlu memperluas topik, membangun akses data operasional teranonimisasi dengan tata kelola privasi, menggunakan metode pembanding yang lebih kuat, dan melaporkan prosedur evaluasi secara lebih transparan.

Kata Kunci: *data mining*, keimigrasian, pemetaan penelitian, PRISMA, tinjauan pustaka sistematis, *Naive Bayes*

A. Pendahuluan

Layanan keimigrasian mengelola mobilitas orang, dokumen perjalanan, izin tinggal, pengawasan orang asing, dan pemeriksaan perbatasan. Digitalisasi proses tersebut menghasilkan data administratif serta data interaksi publik dalam jumlah besar. Data permohonan paspor, pola kedatangan, hasil wawancara, informasi izin tinggal, dan ulasan layanan dapat dianalisis untuk mendukung perencanaan kapasitas, evaluasi pelayanan, serta pengawasan berbasis risiko. Dalam konteks ini, *data mining* digunakan untuk menemukan pola, hubungan, dan kecenderungan yang tidak mudah dikenali melalui pemeriksaan manual.

Penelitian keimigrasian di Indonesia telah menerapkan beragam pendekatan. *K-Means* digunakan untuk mengelompokkan data permohonan paspor dan izin tinggal (Assiroj, Kurnia, et al., 2023; Piranti et al., 2026). Pada sisi pelayanan publik, Naive Bayes, *Support Vector Machine*, *Decision Tree*, *Random Forest*, *Logistic Regression*, dan *Convolutional Long Short-Term Memory* digunakan untuk menganalisis ulasan M-Paspor, Google Maps, dan opini di media sosial (Assiroj, Kurnia, et al., 2023; Elbert & T, 2025; Hariwibowo & Setiawan, 2026; Maheri et al., 2025; Martanto & Istiono, 2024; Pontiselly & Ahsan, 2025). Penelitian lain menggunakan *Latent Dirichlet Allocation* untuk menemukan tema dalam komentar pengguna (Hapsari & Nataliani, 2024), *FP-Growth* untuk memetakan pola Pekerja Migran Indonesia non prosedural (Syahirah, 2025), dan *Backpropagation Neural Network* untuk memprediksi pola pengunjung (Puspitasari, 2018).

Perkembangan tersebut menunjukkan bahwa *data mining* telah digunakan pada beberapa fungsi keimigrasian. Namun, penelitian yang tersedia belum tersebar secara seimbang. Sebagian besar studi

memanfaatkan data publik yang mudah diperoleh, sedangkan data operasional seperti izin tinggal, pola perlintasan, pemeriksaan dokumen, serta indikator pengawasan masih jarang dianalisis. Konsentrasi pada satu jenis data dapat mendorong pengulangan topik dan algoritma, sekaligus membatasi kontribusi penelitian terhadap kebutuhan operasional instansi.

Pemetaan sistematis diperlukan untuk mengetahui topik yang dominan, teknik analisis yang digunakan, algoritma yang dipilih, serta kesenjangan yang masih terbuka. Kajian terdahulu dalam dokumen sumber belum membedakan secara konsisten antara teknik analisis dan bidang *machine learning*. Dalam revisi ini, kategori disusun pada tingkat yang setara, yaitu klasifikasi, klasterisasi, asosiasi, pemodelan topik, dan prediksi. Algoritma kemudian ditempatkan sebagai metode yang menjalankan setiap tugas analitik.

Penelitian ini bertujuan memetakan topik, teknik, dan algoritma *data mining* dalam penelitian keimigrasian. Kajian menjawab tiga pertanyaan: (1) topik apa yang paling banyak dibahas; (2) teknik dan algoritma apa yang dominan; serta (3) kesenjangan apa yang perlu diprioritaskan dalam penelitian berikutnya. Hasilnya diharapkan mendukung penyusunan agenda penelitian yang lebih beragam, relevan, dan bertanggung jawab terhadap perlindungan data.

B. Metode Penelitian

Penelitian menggunakan desain tinjauan pustaka sistematis dengan struktur pelaporan yang mengacu pada PRISMA 2020 (Page et al., 2021). PRISMA digunakan untuk memperjelas sumber informasi, proses identifikasi, kriteria pemilihan, dan penetapan studi yang dianalisis. Pencarian terakhir dilakukan pada 16 Juli 2026 melalui Google Scholar dan OpenAlex.

Strategi Pencarian

Kata kunci awal dalam dokumen sumber adalah "*data mining* keimigrasian". Untuk meningkatkan ketercakupan, pencarian diperluas menggunakan padanan bahasa Indonesia dan Inggris, yaitu "immigration *data mining*", "immigration analytics", "passport clustering", "residence permit clustering", "immigration sentiment analysis", "Indonesian migrant worker *data mining*", dan "border analytics". Kombinasi istilah disesuaikan dengan fasilitas pencarian setiap basis data. Pencarian menghasilkan 259 studi,

yang terdiri atas 200 studi dari Google Scholar dan 59 rekaman dari OpenAlex.

Tabel 1 Hasil pencarian dan seleksi studi

Sumber	Studi ditemukan	Kandidat	Studi terpilih
Google Scholar	200	24	17
OpenAlex	50	5	2
Jumlah	**250**	**29**	**19**

Sumber: Analisis penulis (2026).

Kriteria Seleksi, Ekstraksi, dan Kualitas Pelaporan

Studi dimasukkan apabila berupa artikel jurnal atau prosiding yang membahas penggunaan *data mining* pada layanan, dokumen, migrasi, pengawasan, atau opini publik terkait keimigrasian; menjelaskan sumber data dan metode analisis; serta menyediakan teks lengkap yang dapat ditelaah. Studi dikeluarkan apabila merupakan duplikat, tidak menggunakan metode *data mining*, tidak memiliki hubungan substantif dengan keimigrasian, atau tidak menyajikan informasi yang cukup untuk ekstraksi.

Tabel 2 Kriteria seleksi dan ekstraksi data

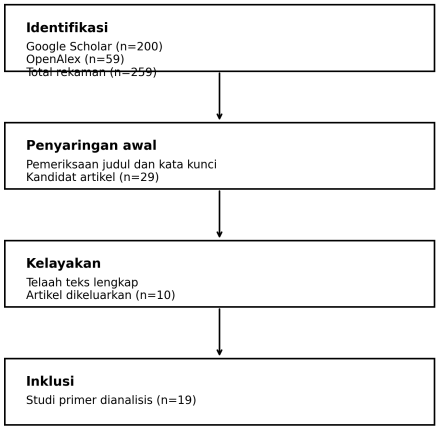
Komponen	Keterangan
Inklusi	Artikel jurnal atau prosiding yang membahas *data mining* pada layanan, pengawasan, dokumen, migrasi, atau opini publik terkait keimigrasian.
Eksklusi	Duplikat, naskah tanpa teks lengkap, studi tanpa metode *data mining*, atau konteks yang tidak berhubungan dengan keimigrasian.
Ekstraksi	Penulis, tahun, topik, jenis data, teknik analisis, algoritma, metrik, dan hasil utama.
Kualitas pelaporan	Kejelasan tujuan, sumber data, metode, prosedur evaluasi, dan keterbatasan.
Sintesis	Tabulasi frekuensi dan analisis naratif untuk mengidentifikasi dominasi serta kesenjangan penelitian.

Sumber: Analisis penulis (2026)

Penyaringan dilakukan melalui pemeriksaan judul dan kata kunci, kemudian telaah teks lengkap. Sebanyak 29 studi dipertahankan sebagai kandidat dan 10 studi dikeluarkan pada tahap kelayakan karena ruang lingkup atau informasi metode tidak memenuhi kriteria. Korpus akhir terdiri atas 19 studi primer. Data diekstraksi ke dalam matriks studi, lalu frekuensi topik, teknik, dan algoritma dihitung. Satu studi dapat dihitung pada lebih dari satu kategori algoritma apabila membandingkan beberapa metode. Oleh karena itu, jumlah frekuensi algoritma dapat melebihi jumlah studi.

Penilaian kualitas dilakukan pada tingkat pelaporan dasar berdasarkan kejelasan tujuan, sumber data, metode, dan hasil evaluasi. Skor rinci per studi tidak tersedia dalam dokumen sumber sehingga penilaian digunakan secara kualitatif dan tidak menghasilkan

peringkat antar studi. Keterbatasan tersebut dilaporkan untuk mencegah kesimpulan yang melebihi bukti.



Gambar 1 Alur identifikasi dan seleksi studi

Gambar 1 menunjukkan penyusutan korpus dari 259 rekaman menjadi 29 kandidat dan 19 studi primer. Log sumber tidak menyimpan jumlah duplikat secara terpisah. Oleh karena itu, diagram menggambarkan alur seleksi yang dapat diverifikasi dari dokumen, bukan rekonstruksi lengkap diagram alir PRISMA.

C. Hasil Penelitian dan Pembahasan

Profil Studi Terpilih

Sebanyak 19 studi terbitan 2018-2026 dianalisis. Studi meningkat pada 2024-2026, terutama pada analisis sentimen layanan digital dan opini publik. Korpus juga mencakup klasterisasi data paspor dan izin tinggal, asosiasi data Pekerja Migran Indonesia, prediksi pola pengunjung, dan penyusunan laporan intelijen untuk pencegahan terorisme. Tabel 3 memetakan topik, teknik analisis, dan algoritma setiap studi.

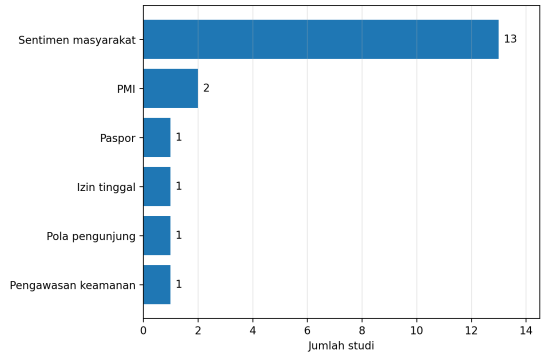
Tabel 3 Pemetaan 19 studi primer

No.	Studi	Topik	Teknik dan algoritma
1	(Piranti et al., 2026)	Paspor	Klasterisasi; *K-Means*
2	(Nur Erlina et al., 2025)	Sentimen masyarakat	Klasifikasi; *Naive Bayes*
3	(Syahirah, 2025)	PMI	Asosiasi; *FP-Growth*
4	(Assiroj, Pribadhi, et al., 2023)	Izin tinggal	Klasterisasi; *K-Means*
5	(Pontiselly & Ahsan, 2025)	Sentimen Masyarakat	Klasifikasi; ConvLSTM
6	(Hapsari & Nataliani, 2024)	Sentimen Masyarakat	Pemodelan topik; LDA
7	(Yuliawati et al., 2024)	Sentimen Masyarakat	Klasifikasi; *Naive Bayes*
8	(Maheri et al., 2025)	Sentimen Masyarakat	Klasifikasi; *Naive Bayes* dan SVM
9	(Martanto & Istiono, 2024)	Sentimen Masyarakat	Klasifikasi; *Naive Bayes*
10	(Ramadhani et al., 2022)	Sentimen Masyarakat	Klasifikasi; *Naive Bayes*
11	(Elbert & T, 2025)	Sentimen Masyarakat	Klasifikasi; SVM dan *Decision Tree*
12	(Malikhatin et al., 2021)	PMI	Klasterisasi; *K-Modes*
13	(Gunawan & Furqan, 2024)	Sentimen masyarakat	Klasifikasi; *Naive Bayes*
14	(Puspitasari, 2018)	Pola pengunjung	Prediksi; *Backpropagation Neural Network*

15	(Mariwy et al., 2025)	Sentimen masyarakat	Klasifikasi; *Naive Bayes*
16	(Assiroj, Kurnia, et al., 2023)	Sentimen masyarakat	Klasifikasi; *Naive Bayes*, SVM dan *Logistic Regression*
17	(Hariwibowo & Setiawan, 2026)	Sentimen masyarakat	Klasifikasi; *Naive Bayes*, *Random Forest*, *Logistic Regression* dan *Decision Tree*
18	(Prasetyo, 2020)	Pengawasan keamanan	Klasifikasi; *Naive Bayes*
19	(Zahra et al., 2024)	Sentimen masyarakat	Klasifikasi; *Naive Bayes*

Sumber: Analisis penulis (2026).

Pemetaan Topik Penelitian



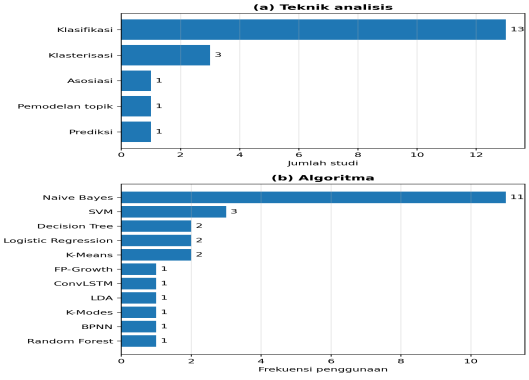
Gambar 2 Distribusi topik penelitian keimigrasian

Gambar 2 menunjukkan bahwa sentimen masyarakat mencakup 13 studi atau 68,4% dari seluruh korpus. Pekerja Migran Indonesia dibahas pada dua studi. Paspor, izin tinggal, pola pengunjung, dan pengawasan keamanan masing-masing hanya muncul satu kali. Koreksi ini berbeda dari pemetaan awal yang memasukkan penelitian pencegahan terorisme ke dalam kategori sentimen. Berdasarkan judul dan tujuan penelitiannya, studi Prasetyo (2020) lebih tepat ditempatkan pada topik pengawasan keamanan.

Dominasi analisis sentimen berkaitan dengan ketersediaan data. Ulasan aplikasi, komentar YouTube, Twitter, TikTok, dan Google Maps dapat diakses tanpa membuka basis data operasional. Data tersebut berguna untuk mengidentifikasi persepsi pengguna, keluhan, dan respons publik. Namun, dominasi data publik membuat penelitian kurang menyentuh kebutuhan lain, seperti pemodelan risiko perjalanan, deteksi anomali izin tinggal, prioritas pemeriksaan, pola pelintas, dan evaluasi akurasi dokumen.

Data paspor dan izin tinggal mengandung data pribadi sehingga aksesnya memerlukan dasar hukum, otorisasi, anonimisasi, pembatasan tujuan, dan pengamanan teknis. Keterbatasan tersebut tidak berarti data operasional tidak dapat digunakan. Penelitian dapat menggunakan data agregat, data sintetis yang tervalidasi, pembelajaran federatif, atau lingkungan pengolahan terkendali untuk mengurangi risiko kebocoran.

Pemetaan Teknik dan Algoritma



Gambar 3 Distribusi teknik analisis dan algoritma

Klasifikasi menjadi teknik utama dengan 13 studi, disusul klasterisasi sebanyak tiga studi. Asosiasi, pemodelan topik, dan prediksi masing-masing muncul satu kali. Kategori *machine learning* tidak lagi diperlakukan sebagai teknik yang setara karena istilah tersebut mencakup beberapa tugas analitik. ConvLSTM diklasifikasikan sebagai algoritma untuk klasifikasi, LDA sebagai pemodelan topik, dan BPNN sebagai prediksi.

Naive Bayes digunakan pada 11 studi. Jumlah ini lebih rendah daripada pemetaan awal karena studi Elbert & T (2025) menggunakan SVM dan *Decision Tree*, bukan *Naive Bayes*. SVM muncul tiga kali. *Decision Tree*, *Logistic Regression*, dan *K-Means* masing-masing dua kali. *FP-Growth*, ConvLSTM, LDA, *K-Modes*, BPNN, dan *Random Forest* masing-masing satu kali. Satu studi dapat menyumbang lebih dari satu frekuensi algoritma ketika melakukan perbandingan model.

Naive Bayes populer karena sederhana, cepat, dan sesuai sebagai *baseline* klasifikasi teks. Akan tetapi, dominasi satu algoritma dapat membatasi kontribusi apabila penelitian hanya melaporkan akurasi tanpa menjelaskan distribusi kelas, ketidakseimbangan data, atau alasan pemilihan fitur. Beberapa studi telah membandingkan *Naive Bayes* dengan SVM, *Logistic Regression*, *Random Forest*, dan *Decision Tree* (Assiroj, Kurnia, et al., 2023; Elbert & T, 2025; Hariwibowo & Setiawan, 2026; Maheri et al., 2025). Perbandingan tersebut perlu menggunakan pembagian data, prapemrosesan, dan metrik yang sama agar hasilnya adil.

Untuk klasifikasi sentimen, akurasi tidak cukup apabila distribusi kelas tidak seimbang. Penelitian sebaiknya melaporkan *precision*, *recall*, *F1-score*, *confusion matrix*, dan AUC sesuai karakteristik masalah. Validasi silang atau pengujian pada periode dan lokasi berbeda juga diperlukan untuk menilai kemampuan generalisasi. Pada klasterisasi, peneliti perlu melaporkan ukuran validitas seperti *silhouette coefficient* atau *Dunn Index*,

sedangkan asosiasi perlu menjelaskan *support*, *confidence*, dan *lift*.

Kesenjangan dan Agenda Penelitian

Tabel 4 Kesenjangan dan rekomendasi penelitian

Kesenjangan	Rekomendasi
Topik terkonsentrasi pada sentimen	Mengembangkan studi izin tinggal, pola kedatangan, risiko perjalanan, pengawasan perbatasan dan deteksi anomali.
Penggunaan data operasional terbatas	Membangun skema akses data teranonimisasi dengan persetujuan instansi, pembatasan tujuan dan pengendalian privasi.
Algoritma kurang beragam	Menguji *ensemble, deep learning, transformer,* deteksi anomali dan *explainable artificial intelligence* dengan *baseline* yang adil.
Evaluasi belum seragam	Melaporkan pembagian data, prapemrosesan, *baseline, precision, recall, F1-score,* AUC dan validasi lintas lokasi atau waktu.
Konteks dominan domestik dari satu lokasi	Mendorong penelitian antar kantor imigrasi dan lintas negara dengan definisi variabel serta protokol yang setara.
Pelaporan kualitas belum lengkap	Menyimpan log pencarian, keputusan penyaringan, karakteristik dataset, kode analisis dan keterbatasan setiap studi.

Sumber: Analisis penulis (2026).

Tabel 4 menunjukkan bahwa agenda penelitian tidak cukup diarahkan pada penambahan algoritma. Akses data, kualitas pelaporan, desain evaluasi, perlindungan privasi, dan hubungan hasil model dengan keputusan keimigrasian harus menjadi bagian dari rancangan. Penggunaan model yang lebih kompleks tidak otomatis menghasilkan kontribusi yang lebih tinggi apabila data tidak representatif atau evaluasi tidak dapat di replikasi.

Pada tataran praktis, *data mining* dapat mendukung perencanaan kapasitas layanan, identifikasi pola permohonan, evaluasi kanal digital, pemetaan pola PMI non prosedural, dan penyusunan indikator pengawasan. Hasil model perlu diterjemahkan menjadi keluaran yang

dapat ditindaklanjuti. Analisis sentimen, misalnya, akan lebih bernilai apabila dikaitkan dengan jenis keluhan, versi aplikasi, lokasi kantor, atau waktu penyelesaian layanan. Klaster permohonan paspor akan lebih berguna apabila diterjemahkan menjadi kebutuhan loket, jadwal layanan, atau pola wawancara.

Pemanfaatan analitik untuk pengawasan memerlukan kehati-hatian lebih tinggi karena kesalahan model dapat memengaruhi hak pelintas. Model harus digunakan sebagai dukungan keputusan, bukan satu-satunya dasar tindakan. Peneliti perlu menjelaskan sumber data, variabel yang digunakan, mekanisme koreksi, potensi bias, dan peran petugas dalam meninjau hasil. Prinsip minimalisasi data dan pembatasan tujuan harus diterapkan sejak perancangan penelitian.

Keterbatasan Penelitian

Kajian ini memiliki beberapa keterbatasan. Pertama, pencarian hanya menggunakan Google Scholar dan OpenAlex sehingga studi yang terindeks di Scopus, Web of Science, IEEE Xplore, atau basis data nasional lain dapat terlewat. Kedua, kata kunci awal lebih berorientasi pada istilah keimigrasian Indonesia. Ketiga, dokumen sumber tidak menyimpan jumlah duplikat dan alasan eksklusi per artikel secara rinci sehingga diagram PRISMA lengkap tidak dapat direkonstruksi. Keempat, beberapa metadata studi pada versi awal tidak lengkap dan telah diperbaiki sejauh dapat diverifikasi, tetapi penilaian risiko bias pada setiap studi belum tersedia. Kelima, meta-analisis tidak dilakukan karena tujuan, dataset, dan metrik penelitian sangat beragam.

D. Kesimpulan

Tinjauan terhadap 19 studi menunjukkan bahwa penelitian *data mining* keimigrasian masih terkonsentrasi pada sentimen masyarakat, yaitu 13 studi. Klasifikasi merupakan teknik paling dominan dengan 13 studi, sedangkan *Naive Bayes* merupakan algoritma yang paling sering digunakan dengan 11 penerapan. Penelitian mengenai PMI muncul dua kali, sementara paspor, izin tinggal, pola pengunjung, dan pengawasan keamanan masing-masing hanya satu kali. Temuan tersebut menunjukkan bahwa perkembangan penelitian lebih banyak mengikuti ketersediaan data publik daripada keragaman kebutuhan operasional keimigrasian. Penelitian

berikutnya perlu memperluas topik, membangun mekanisme akses data teranonimisasi yang sah, membandingkan algoritma dengan protokol yang setara, menggunakan metrik yang sesuai, dan melakukan validasi lintas lokasi atau waktu. Pembaruan SLR juga perlu menyimpan log pencarian, duplikasi, alasan eksklusi, dan penilaian kualitas per studi agar proses kajian dapat direplikasi dan diaudit.

DAFTAR PUSTAKA

Assiroj, P., Kurnia, A., & Alam, S. (2023). The performance of Naïve Bayes, support vector machine, and logistic regression on Indonesia immigration sentiment analysis. *Bulletin of Electrical Engineering and Informatics*, *12*(6), 3843–3852. https://doi.org/10.11591/eei.v12i6.5688

Assiroj, P., Pribadhi, I. A., & Rizal S, M. (2023). A K-MEANS ALGORITHM FOR THE IMMIGRATION RESIDENTIAL PERMIT. *Jurnal Ilmiah Kajian Keimigrasian*, *6*(2), 145–152. https://doi.org/10.52617/jikk.v6i2.468

Elbert, & T, T. (2025). Analyzing User Perception of The M-Paspor Application through Aspect-based Sentiment Analysis with SVM and Decision Tree. *2025 5th International Conference of Science and Information Technology in Smart Administration (ICSINTESA)*, 457–462. https://doi.org/10.1109/ICSINTESA68165.2025.11413789

Gunawan, I., & Furqan, Mhd. (2024). *Analisis Sentimen Terhadap Tindakan Pemerintah Indonesia Untuk Menampung Sementara Pengungsi Etnis Rohingya Menggunakan Naïve Bayes*.

Hapsari, T. S., & Nataliani, Y. (2024). TEXT MINING WITH LATENT DIRICHLET ALLOCATION FOR ANALYZING PUBLIC COMMENTS ON THE M-PASSPORT APPLICATION. *Jurnal Teknik Informatika (Jutif)*, *5*(4), 1093–1101. https://doi.org/10.52436/1.jutif.2024.5.4.1929

Hariwibowo, I. N., & Setiawan, W. Y. (2026). Evaluation of the M-Paspor App as a Public Service: A Sentiment Analysis. *Jurnal Manajemen Pelayanan Publik*, *10*(1), 206–219. https://doi.org/10.24198/jmpp.v10i1.68645

Maheri, R., Salisah, F. N., Muttakin, F., & Megawati, M. (2025). ANALISIS SENTIMEN ULASAN APLIKASI M-PASPOR MENGGUNAKAN NAIVE BAYES DAN SUPPORT VECTOR MACHINE. *JIPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, *10*(1), 448–458.

https://doi.org/10.29100/jipi.v10i1.5826

Malikhatin, H., Rusgiyono, A., & Maruddani, D. A. I. (2021). PENERAPAN k-MODES CLUSTERING DENGAN VALIDASI DUNN INDEX PADA PENGELOMPOKAN KARAKTERISTIK CALON TKI MENGGUNAKAN R-GUI. *Jurnal Gaussian*, *10*(3), 359–366. https://doi.org/10.14710/j.gauss.v10i3.32790

Mariwy, F., Watratan, C. J. G., Santos, R. Dos, Reba, D. N., Sabono, R. F., & Pasiakan, N. I. (2025). Analisis Sentimen Komentar Tiktok Tentang Kasus TPPO TKI Di Kamboja Menggunakan Klasifikasi Naive Bayes Untuk Mengukur Opini Publik. *Techno.Com*, *24*(2), 602–616. https://doi.org/10.62411/tc.v24i2.12793

Martanto, M. L. F., & Istiono, W. (2024). Sentiment Analysis of M-Paspor App Reviews Using Multinomial Naive Bayes. *Journal of Logistics, Informatics and Service Science*, *11*(10). https://doi.org/10.33168/JLISS.2024.1017

Nur Erlina, N., Trinata, C., & Anung Prabadhi, I. (2025). ANALISIS SENTIMEN ULASAN KEIMIGRASIAN PADA GOOLE MAPS MENGGUNAKAN ALGORITMA NAIVE BAYES. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *9*(4), 6613–6618. https://doi.org/10.36040/jati.v9i4.13820

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Piranti, N. M., Khalish, F., & Martadireja, O. P. (2026). Passport Application Data Mining Implementation at The Muara Enim Non-TPI Class II Immigration Office. *Inkubis : Jurnal Ekonomi Dan Bisnis*, *8*(1), 42–62. https://doi.org/10.59261/inkubis.v8i1.141

Pontiselly, P., & Ahsan, M. (2025). Monitoring Sentimen Ulasan Masyarakat di Google Maps Terhadap Layanan Kantor Imigrasi Surabaya Sebagai Early Warning System untuk Perbaikan Kualitas Berbasis Convolutional Long Short Term Memory. *Jurnal Sains Dan Seni ITS*, *13*(2). https://doi.org/10.12962/j23373520.v13i2.145412

Prasetyo, Y. E. B. (2020). DESIGN OF DATA MINING METHOD IN MAKING INTELLIGENT RESULTS REPORT TO PREVENT CRIMINAL ACTS OF

TERRORISM IN THE SEMARANG IMMIGRATION OFFICE. *TEMATICS: Technology ManagemenT and Informatics Research Journals*, *2*(2), 67–81. https://doi.org/10.52617/tematics.v2i2.105

Puspitasari, I. A. A. (2018). Prediksi Pola Pengunjung untuk Meningkatkan Kualitas Layanan Menggunakan Teknik Heuristik Pada Optimasi Bobot Backpropagation Neural Network. *Joined Journal (Journal of Informatics Education)*, *1*(1), 1. https://doi.org/10.31331/joined.v1i1.607

Ramadhani, R. P., Bustami, B., & Razi, A. (2022). KLASIFIKASI TINGKAT KEPUASAN MASYARAKAT DALAM KUALITAS PELAYANAN PEMBUATAN PASPOR DI MASA PANDEMI COVID-19 MENGGUNAKAN ALGORITMA NAIVE BAYES. *Jurnal Teknologi Terapan and Sains 4.0*, *3*(1), 677. https://doi.org/10.29103/tts.v3i1.6944

Syahirah, D. (2025). Implementation of Association Rule Mining Using the FP-Growth Algorithm on Non-Procedural Indonesian Migrant Worker (PMI) Data in South Sulawesi. In *PASUNDAN SOCIAL SCIENCE DEVELOPMENT* (Vol. 6, Number 1).

Yuliawati, Y., Sulistyowati, N., & Siska, S. (2024). OPINION MINING TERHADAP PENANGANAN ROHINGYA DI PROVINSI ACEH MENGGUNAKAN NAÏVE BAYES PADA MEDIA SOSIAL TWITTER. *Jurnal Informatika Dan Teknik Elektro Terapan*, *12*(3). https://doi.org/10.23960/jitet.v12i3.4990

Zahra, A., Mayasari, R., & Pernamasari, I. (2024). Analisis Sentimen pada Aplikasi M-Paspor Menggunakan Algoritma Naïve Bayes Classifier. *Action Research Literate*, *8*(8), 2365–2371. https://doi.org/10.46799/arl.v8i8.466