
Analisis Setimen Tentang Undang-Undang Perlindungan Data Pribadi Menggunakan Algoritma *Naïve Bayes Classifier*

Melin Manipi¹, Rendra Soekarta², Muhammad Yusuf³, Fitriyani Tella⁴

^{1,2,3}Program Studi Teknik Informatika, Universitas Muhammadiyah Sorong

e-mail: ¹melinmanipi25@gmail.com, ²rendrasoekarta@gmail.com, ³yusuf@um-sorong.ac.id,

⁴fitriyanitella@um-sorong.ac.id

Abstrak

Undang-undang Perlindungan Data Pribadi merupakan suatu peraturan yang telah resmi disahkan oleh pemerintah Indonesia pada tanggal 10 September 2022. Peraturan ini disahkan dalam rangka melindungi dan menjamin hak warga negara Indonesia terkait kerahasiaan informasi dan komunikasi pribadi serta hak atas privasi dan data pribadi warga negara tersebut. Pengesahan tersebut menuai tanggapan pro dan kontra dari kalangan masyarakat Indonesia. Tanggapan tersebut banyak dituangkan pada media sosial berita resmi yang memberitakan terkait pengesahan Undang-undang tersebut. Tanggapan pro yang dituangkan seperti masyarakat mendukung pengesahan Undang-undang tersebut karena menganggap hal tersebut menjadi salah satu solusi yang efektif untuk kasus penyalahgunaan dan kebocoran data sedangkan, kontranya seperti masyarakat menganggap pengesahan tersebut hanyalah formalitas saja. Berdasarkan hal tersebutlah, diperlukan sebuah solusi yaitu penerapan bidang data mining analisis sentimen yang dapat membantu dalam menganalisis tanggapan masyarakat sehingga dapat diketahui tingkat dominasi tanggapannya. Pengumpulan tanggapan terkait Undang-undang Perlindungan Data Pribadi diambil dari platform Twitter dan Youtube dengan perolehan data yaitu sebesar 744. Analisis sentimen dilakukan dengan menerapkan metode *Naïve Bayes Classifier* sebagai algoritma klasifikasi tanggapan. Penelitian ini menghasilkan sebuah sistem analisis sentimen dengan perolehan nilai akurasi sebesar 90,604%. Tujuan eksternal dari penelitian ini ialah mendapatkan bahan evaluasi bagi pemerintah Indonesia terkait penerapan kebijakan Undang-undang Perlindungan Data Pribadi.

Kata kunci : Analisis Sentimen, *Naïve Bayes Classifier*, Undang-undang Data Pribadi

1. PENDAHULUAN

Penggunaan internet yang semakin meningkat membutuhkan keamanan digital, sementara saat ini terdapat banyak penyalahgunaan data pribadi dalam ranah digital untuk aksi kejahatan cyber. Kebocoran data pribadi secara digital juga mengancam para pengguna internet untuk itu pengesahan Rancangan Undang-undang Perlindungan Data Pribadi menjadi penting agar menjamin keamanan ekosistem digital saat ini. Rancangan Undang-Undang Perlindungan Data Pribadi atau RUU PDP adalah peraturan yang telah disahkan oleh pemerintah Indonesia pada tanggal 10 September 2022 sebagai landasan bagi penegak hukum guna melindungi dan menjamin hak warga Negara Indonesia yang mengatur tentang kerahasiaan informasi dan komunikasi pribadi serta perlindungan hak atas privasi dan data pribadi [1].

Pengesahan RUU PDP menuai tanggapan pro dan kontra dari kalangan masyarakat. Tanggapan pronya seperti masyarakat mendukung pengesahan RUU PDP karena dengan adanya RUU PDP dianggap menjadi salah satu solusi yang efektif dalam kasus kebocoran atau

penyalahgunaan data pribadi. sedangkan kontranya adalah RUU PDP dianggap tidak detail mengenai jenis-jenis data pribadi yang masuk dalam kualifikasi sehingga masyarakat beranggapan bahwa RUU PDP hanya disahkan sebagai formalitas dan data pribadi yang seharusnya dilindungi berdasarkan RUU PDP justru dapat disalahgunakan oleh pemerintah [2].

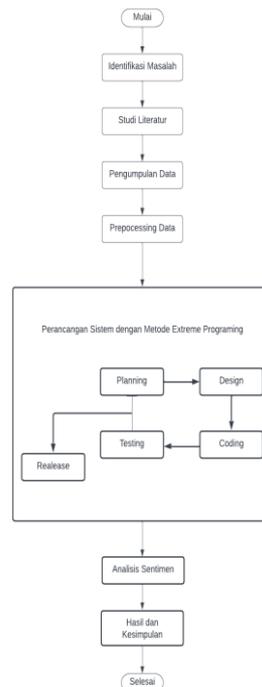
Analisis sentimen adalah suatu proses data mining digunakan untuk mengidentifikasi suatu pendapat atau opini seseorang terhadap suatu isu atau topik tertentu [3]. Dengan memanfaatkan bidang analisis sentimen, dapat diketahui tingkat dominasi opini atau tanggapan pro dan kontra terkait suatu topik. Penggunaan bidang analisis sentimen juga memerlukan sebuah algoritma klasifikasi, yang berperan sebagai algoritma pengklasifikasian tanggapan tersebut, salah satunya adalah *Naïve Bayes Classifier*. Algoritma *Naïve Bayes Classifier* merupakan algoritma *machine learning* yang termasuk dalam *supervised classification* [4]. Algoritma ini juga banyak digunakan untuk kasus klasifikasi teks diantaranya seperti menyaring spam email, analisis sentimen, sistem rekomendasi, dsb [4].

Beberapa penelitian telah menerapkan analisis sentimen dengan metode *Naïve Bayes Classifier* untuk mendapatkan tingkat dominasi opini atau tanggapan terkait suatu topik diantaranya, analisis sentimen terhadap bakal calon presiden Indonesia 2024 [5], analisis sentimen tanggapan masyarakat cyberbullying di media sosial, analisis sentimen physical distancing pada twitter [6], analisis sentimen dokumen twitter mengenai dampak virus corona [7], analisis sentimen terhadap review restoran di Singapura [8] dan analisis sentimen terhadap pembelajaran daring [9]. Penelitian-penelitian tersebut menunjukkan bahwa analisis sentimen adalah suatu bidang yang dapat diterapkan dan digunakan sebagai solusi untuk membantu manusia dalam hal mendapatkan tingkat dominasi suatu opini atau tanggapan.

Berdasarkan uraian tersebut, perlunya dilakukan penelitian dengan pembahasan analisis sentimen terkait Undang-undang Perlindungan Data Pribadi.

2. METODE PENELITIAN

2.1 Alur Penelitian



Gambar 2 *Dataset*2.2.4 *Preprocessing Data*

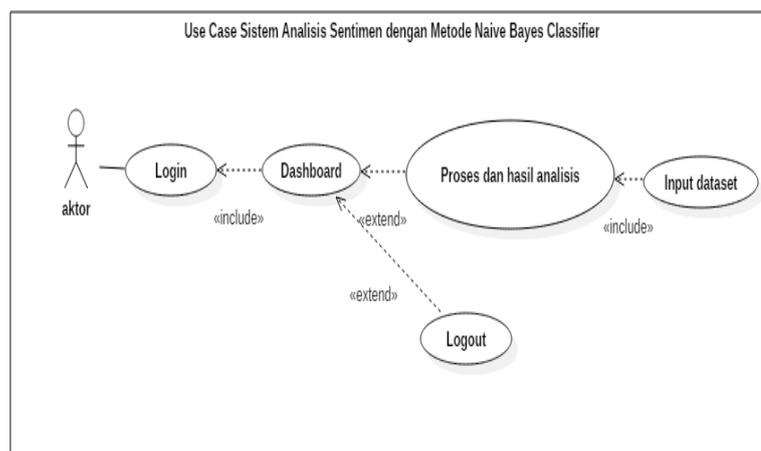
Preprocessing data dilakukan terhadap dataset yang telah diperoleh pada tahap sebelumnya. *Preprocessing* data adalah proses pembersihan dan pengoptimalan data agar dapat mencapai hasil yang maksimal [3]. Terdapat beberapa metode *Preprocessing* data yang digunakan diantaranya yaitu [3]:

- 1) *Stemming*
Metode untuk mengubah suatu data teks yang berupa kata menjadi kata dasarnya seperti menjahit jadi jahit.
- 2) *Casefolding*
Metode untuk mengubah seluruh huruf yang terdapat pada data teks menjadi huruf dengan tipe atau jenis *lowercase* (huruf kecil).
- 3) *Cleansing*
Metode untuk menghapus karakter yang tidak berguna pada data teks seperti titik, koma, tanda tanya, dsb.
- 4) *Tokenization*
Metode untuk mengubah kalimat yang terdapat pada data teks menjadi terpisah setiap katanya.
- 5) *Stopword*
Metode untuk menghilangkan kata yang tidak bermakna yang terdapat pada data teks.
- 6) *Removal*
Metode untuk memfilter kata-kata yang penting yang terdapat pada data teks.

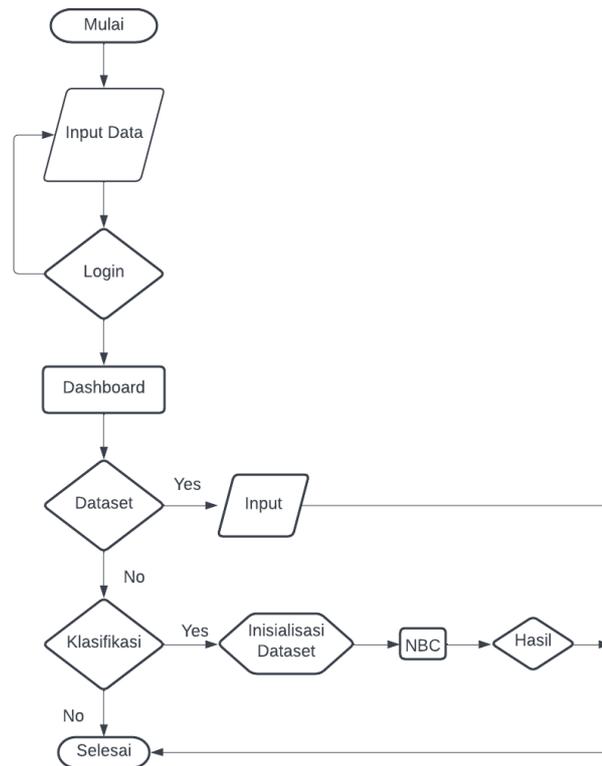
2.2.5 *Perancangan Sistem*

Perancangan sistem dilakukan dengan menerapkan metode pengembangan perangkat lunak *Extreme Programming (XP)*. Metode *XP* ini terbagi atas beberapa langkah yaitu :

- 1) *Planning*
Langkah perencanaan, merencanakan kebutuhan-kebutuhan yang diperlukan dalam proses perancangan dan pengembangan sistem.
- 2) *Design*
Langkah perancangan desain sistem. Rancangan dibuat untuk merepresentasikan sistem yang akan dikembangkan. Rancangan sistem dibuat menggunakan *Use Case* dan *Flowchart* diagram. *Use Case* adalah diagram yang menyatakan visualisasi interaksi yang terjadi antara pengguna sistem dan sistem[10]. *Flowchart* adalah diagram yang menyatakan representasi grafis dari urutan atau langkah-langkah dari suatu sistem atau program atau prosedur [11]. Berikut terdapat gambar dari *Use Case* dan *Flowchart* sistem :



Gambar 3 Use Case Sistem



Gambar 4 Flowchart Sistem

Selain merancang sistem dengan *Use Case* dan *Flowchart* diagram, dibuat pula pemodelan dari analisis sentimen dengan metode *Naïve Bayes Classifier*. Pemodelan *Naïve Bayes Classifier* dibangun sebagai acuan dalam pembuatan kode program sistem.

3) *Coding*

Langkah *coding*, melakukan implementasi dalam bentuk kode program berdasarkan rancangan sistem yang telah dibuat pada langkah sebelumnya. Bahasa pemrograman yang digunakan adalah *PHP*. Implementasi dalam bentuk kode program hingga sistem berhasil terbangun secara keseluruhan.

4) *Testing*

Langkah *testing*, melakukan pengujian terhadap sistem yang telah berhasil dibangun. Metode pengujian sistem yang digunakan ialah *Black Box Testing*. Metode *Black Box* merupakan metode pengujian yang dilakukan dengan cara tanpa memeriksa detail sistem namun, hanya memeriksa hasil dari nilai inputan dan nilai output.

2.2.6 Analisis Sentimen

Analisis sentimen dilakukan menggunakan dataset yang telah berhasil dikumpulkan dan menggunakan sistem yang telah berhasil dibangun. Analisis sentimen dilakukan untuk mengetahui tingkat dominasi tanggapan pro dan kontra (positif dan negatif) terkait Undang-undang Perlindungan Data Pribadi.

2.2.7 Hasil dan Kesimpulan

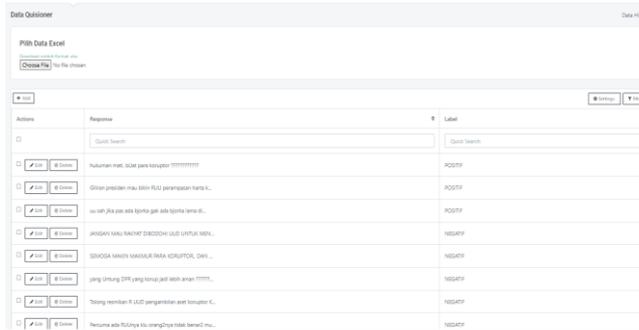
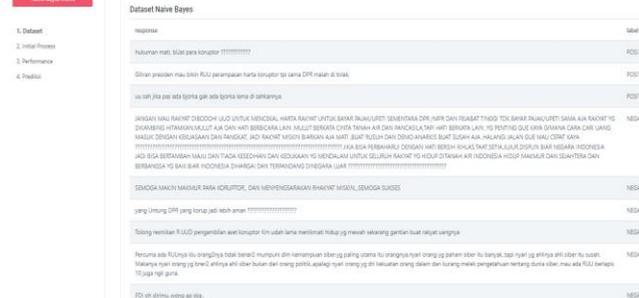
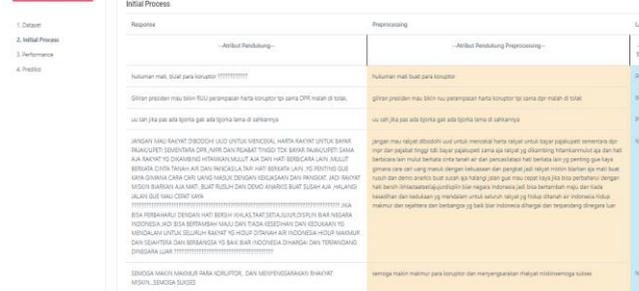
Pada tahap ini dilakukan penyusunan hasil penelitian dalam bentuk paper (karya ilmiah).

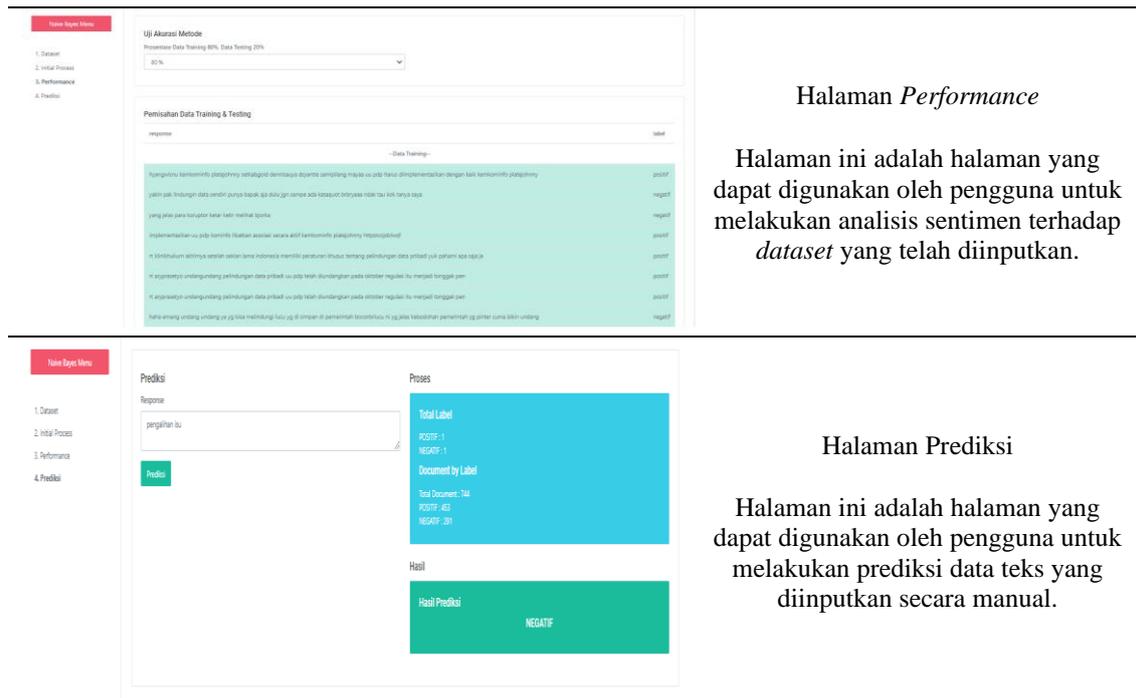
3. HASIL DAN PEMBAHASAN

3.1 Implementasi Interface Sistem

Berikut merupakan implementasi *interface* atau tampilan sistem pada aplikasi analisis sentimen menggunakan metode *Naïve Bayes Classifier*.

Tabel 1 Implementasi Interface Sistem

Gambar Antarmuka	Keterangan
	<p>Halaman <i>Dashboard</i> Sistem</p> <p>Halaman <i>dashboard</i> adalah halaman utama dari sistem yang tampil pada saat pengguna mengakses sistem.</p>
	<p>Halaman <i>Dataset</i></p> <p>Halaman ini adalah halaman yang dapat digunakan oleh pengguna sistem untuk mengelola (input, edit, hapus) <i>dataset</i>.</p>
	<p>Halaman <i>Show Dataset</i></p> <p>Halaman ini adalah halaman yang dapat digunakan oleh pengguna untuk melihat <i>dataset</i> yang telah diinputkan ke dalam sistem.</p>
	<p>Halaman <i>Inisialisasi</i></p> <p>Halaman ini adalah halaman yang digunakan oleh pengguna untuk dapat melakukan penginisialisasian atau pengenalan <i>dataset</i>. Pengenalan dalam hal ini adalah mengenali konten/komentar dan label/kelas klasifikasi.</p>



Halaman *Performance*

Halaman ini adalah halaman yang dapat digunakan oleh pengguna untuk melakukan analisis sentimen terhadap *dataset* yang telah diinputkan.

Halaman *Prediksi*

Halaman ini adalah halaman yang dapat digunakan oleh pengguna untuk melakukan prediksi data teks yang diinputkan secara manual.

3.2 Hasil Analisis Sentimen Undang-undang Perlindungan Data Pribadi

Berdasarkan jumlah *dataset* yang telah diperoleh yaitu 744 data dengan menerapkan metode *splitting dataset* 80:20 (data latih:data uji) memperoleh hasil yaitu sebagai berikut :

True Positive	True Negative	False Positive	False Negative
53	82	8	6

Hasil Akurasi : 90.604%
Hasil Presisi : 86.885%
Hasil Recall : 89.831%

Gambar 5 Hasil Analisis Sentimen Undang-undang Perlindungan Data Pribadi

4. KESIMPULAN

Berdasarkan pembahasan di atas, dapat disimpulkan beberapa kesimpulan meliputi :

1. Perancangan dan pembangunan sistem analisis sentimen dengan algoritma *Naïve Bayes Classifier* telah berhasil dilakukan yang dibuktikan dengan terbangunnya suatu sistem tersebut dan terbukti berhasil dalam menentukan tingkat dominasi opini atau tanggapan dari Undang-undang Perlindungan Data Pribadi.
2. Hasil dari analisis sentimen adalah memperoleh nilai akurasi sebesar 90,604%, recall sebesar 89,831% dan presisi sebesar 86,885% berdasarkan *dataset* dan metode *splitting dataset* 80:20. Diketahui pula tingkat dominasi opini atau tanggapan masyarakat tentang Undang-undang Perlindungan Data Pribadi adalah negatif.

5. SARAN

Pada penelitian selanjutnya disarankan dapat melakukan analisis sentimen dengan topik lainnya dan dapat menggunakan metode klasifikasi yang lain seperti *Support Vector Machine* atau dapat menerapkan metode komparasi.

DAFTAR PUSTAKA

- [1] K. T. V Youtube, “Urgensi RUU PDP.” 2022, [Daring]. Tersedia pada: <https://youtu.be/tmlUV7rwbjY>.
 - [2] M. H. Hisbulloh, “Urgensi Rancangan Undang-Undang (RUU) Perlindungan Data Pribadi,” *J. Huk.*, vol. 37, no. 2, hal. 119–133, 2021, doi: 10.26532/jh.v37i2.16272.
 - [3] N. L. P. Merawati, A. Z. Amrullah, dan Ismarmiaty, “Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 1, hal. 123–131, 2021, doi: 10.29207/resti.v5i1.2587.
 - [4] A. Perdana, A. Hermawan, dan D. Avianto, “Analisis Sentimen Terhadap Isu Penundaan Pemilu di Twitter Menggunakan Naive Bayes Clasifier,” *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 11, no. 2, hal. 195–200, 2022, doi: 10.32736/sisfokom.v11i2.1412.
 - [5] M. R. F. Sya’bani, U. Enri, dan T. N. Padilah, “Analisis Sentimen Terhadap Bakal Calon Presiden 2024 dengan Algoritma Naive Bayes,” *J. Ris. Komputer*, vol. 9, no. 2, hal. 265–273, 2022, doi: 10.30865/jurikom.v9i2.3989.
 - [6] H. S. Rifai, S. Febrianti, dan I. Santoso, “ANALISIS SENTIMEN TANGGAPAN MASYARAKAT TERHADAP CYBERBULLYING DI MEDIA SOSIAL MENGGUNAKAN ALGORITMA NAIVE BAYES (NB),” vol. 7, no. 2, hal. 183–196, 2023.
 - [7] N. M. A. J. Astari, Dewa Gede Hendra Divayana, dan Gede Indrawan, “Analisis Sentimen Dokumen Twitter Mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier,” *J. Sist. dan Inform.*, vol. 15, no. 1, hal. 27–29, 2020, doi: 10.30864/jsi.v15i1.332.
 - [8] V. A. Permadi, “Analisis Sentimen Menggunakan Algoritma Naive Bayes Terhadap Review Restoran di Singapura,” *J. Buana Inform.*, vol. 11, no. 2, hal. 141–151, 2020, doi: 10.24002/jbi.v11i2.3769.
 - [9] D. Musfiroh, U. Khaira, P. E. P. Utomo, dan T. Suratno, “Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, hal. 24–33, 2021, doi: 10.57152/malcom.v1i1.20.
 - [10] H. Kurniawan, W. Apriliah, I. Kurnia, dan D. Firmansyah, “Penerapan Metode Waterfall Dalam Perancangan Sistem Informasi Penggajian Pada Smk Bina Karya Karawang,” *J. Interkom J. Publ. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 14, no. 4, hal. 13–23, 2021, doi: 10.35969/interkom.v14i4.78.
 - [11] I. A. Ridlo, “Pedoman Pembuatan Flowchart,” in *Academia.Edu*, Departemen Administrasi Dan Kebijakan Kesehatan, 2017, hal. 27.
-