BAB II
TINJAUAN LITERATUR

2.1 Board of Peace

Board of Peace (BoP) merupakan sebuah forum internasional yang diinisiasi dengan tujuan utama mendorong resolusi konflik dan menjaga stabilitas perdamaian global. Kendati memiliki tujuan yang positif, wacana keikutsertaan Indonesia sebagai anggota dalam organisasi tersebut telah memicu perdebatan publik dan polarisasi opini yang masif, terutama di ruang diskusi digital seperti platform X (sebelumnya Twitter).

Dinamika pro dan kontra yang berkembang berkaitan erat dengan identitas sosiopolitik Indonesia sebagai negara berpenduduk mayoritas Muslim, yang secara historis maupun konstitusional memiliki komitmen teguh dalam mendukung kemerdekaan Palestina. Dari perspektif publik yang menolak (sentimen negatif), sorotan utama ditujukan pada keterlibatan Israel di dalam struktur kelembagaan BoP. Partisipasi Indonesia di dalam forum tersebut dikhawatirkan dapat mencederai nilai-nilai solidaritas serta melemahkan posisi Indonesia terhadap perjuangan kemerdekaan rakyat Palestina.

Sebaliknya, dari perspektif publik yang mendukung (sentimen positif), keanggotaan ini dipandang sebagai manuver diplomasi yang strategis. Kelompok ini menilai bahwa dengan bergabungnya Indonesia secara formal ke dalam BoP, pemerintah akan memiliki ruang diplomasi yang lebih inklusif dan proaktif. Melalui keterlibatan langsung di dalam sistem tersebut, Indonesia diharapkan dapat lebih efektif dalam menyuarakan resolusi perdamaian di

Gaza dan mengawal distribusi bantuan kemanusiaan pada tingkat internasional.

2.2 Analisis Sentimen

Analisis sentimen merupakan cabang dari text mining yang bertujuan untuk mengidentifikasi, mengekstraksi, dan mengklasifikasikan opini publik ke dalam sentimen tertentu, umumnya positif, negatif, atau netral. Dalam ranah kebijakan publik dan diplomasi internasional, analisis ini menjadi instrumen krusial untuk memetakan respons masyarakat terhadap keputusan strategis pemerintah. Penelitian oleh Mao (2024) menunjukkan bahwa pemanfaatan *Natural Language Processing* (NLP) sangat efektif dalam menangkap dinamika opini publik secara komputasional. Lebih lanjut, Leandro & Fianty (2025) menegaskan bahwa analisis sentimen memberikan gambaran yang presisi mengenai isu sosial-politik yang berkembang cepat di media daring, seperti isu keanggotaan Indonesia di Board of Peace.

Analisis sentimen juga telah banyak diterapkan pada berbagai isu kebijakan publik yang berkembang di media sosial. Penelitian Ibrahim dkk., (2025) mengenai sentimen publik terhadap kebijakan RUU TNI menunjukkan bahwa algoritma *Support Vector Machine* mampu mengklasifikasikan opini masyarakat ke dalam kelas positif, netral, dan negatif dengan tingkat akurasi yang tinggi, sehingga metode ini dinilai efektif untuk menganalisis persepsi masyarakat terhadap suatu kebijakan pemerintah.

2.3 Media Sosial sebagai Sumber Data

Platform X (sebelumnya Twitter) merupakan salah satu media sosial berbasis microblogging terbesar yang menyediakan data teks melimpah dalam bentuk tweet. Data dari platform ini mencerminkan opini publik yang spontan, terbuka, dan aktual terhadap suatu isu kebijakan nasional maupun internasional secara real-time.

Penelitian oleh Hidayati & Yamasari, (2025) menunjukkan bahwa platform X merupakan media microblogging yang banyak digunakan masyarakat untuk menyampaikan opini secara spontan terhadap berbagai isu sosial sehingga menjadi sumber data yang sangat potensial dalam analisis sentimen.

Dalam konteks kebijakan luar negeri, respons publik yang muncul di platform X mengenai keanggotaan Indonesia di Board of Peace menunjukkan adanya variasi dan polarisasi opini yang luas. Oleh karena itu, dokumen teks dari platform X ini menjadi sumber data primer yang sangat valid untuk dibedah pola sentimennya menggunakan pendekatan pembelajaran mesin.

2.4 Preprocessing dan TF-IDF

Kualitas model klasifikasi sangat bergantung pada tahapan pra-pemrosesan (*preprocessing*) untuk mengubah data teks tidak terstruktur menjadi format yang siap diolah. Tahapan ini meliputi *cleaning, case folding, tokenizing, stopword removal,* dan *stemming* (Syahputra dkk., 2022). Setelah data teks bersih, dilakukan proses pembobotan kata menggunakan metode *Term Frequency - Inverse Document Frequency* (TF-IDF) untuk mengekstraksi fitur teks secara optimal (Lestari & Hutagalung, 2025). Metode ini memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam sebuah

dokumen dan kelangkaannya dalam seluruh korpus. Rumus matematis TF-IDF adalah sebagai berikut:

$$W_{t,d} = \mathrm{T}F_{t,d} \times \log\left(\frac{N}{df_t}\right)$$

Keterangan:

1. $W_{t,d}$*: Bobot kata* t *pada dokumen* d.
2. $TF_{t,d}$ *: Frekuensi kemunculan kata* t *dalam dokumen* d
3. N *: Total seluruh dokumen dalam dataset.*
4. df_t *: Jumlah dokumen yang mengandung kata* t.

2.5 Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah algoritma pembelajaran mesin (*machine learning*) yang bekerja dengan mencari *hyperplane* atau bidang pemisah optimal untuk membagi dua kelas sentimen dalam ruang fitur berdimensi tinggi. SVM dikenal unggul dalam menangani data teks yang bersifat *sparse* dan tahan terhadap risiko *overfitting* dalam klasifikasi opini media sosial (Leandro & Fianty, 2025). Persamaan dasar *hyperplane* pada SVM untuk memisahkan kelas didefinisikan sebagai:

$$\mathrm{f}(x) = \mathrm{sign}(w \cdot x + b)$$

Keterangan:

1. w : Vektor bobot (*weight vector*) yang menentukan arah posisi *hyperplane*.
2. x : Representasi fitur dari data teks (vektor input).
3. b : Nilai *bias* yang menentukan posisi pergeseran *hyperplane* dari titik koordinat pusat.
4. sign : Fungsi tanda (*signum function*) yang menentukan kelas akhir berdasarkan nilai hasil perhitungan (positif atau negatif).

Dalam praktiknya pada klasifikasi teks, data sentimen seringkali tidak dapat dipisahkan secara linier dengan sempurna. Untuk mengatasi hal ini, SVM menggunakan pendekatan margin lembut (*soft-margin*) dengan mengandalkan parameter penalti C. Parameter C ini berfungsi sebagai pengontrol toleransi kesalahan klasifikasi. Nilai C digunakan algoritma untuk mencari titik keseimbangan terbaik antara membentuk margin pemisah yang paling lebar dengan meminimalkan jumlah kesalahan prediksi pada data latih:

2.6 Naïve Bayes

Naïve Bayes merupakan algoritma klasifikasi berbasis probabilitas yang menerapkan Teorema Bayes dengan asumsi independensi antar fitur (setiap kata dalam kalimat dianggap berdiri sendiri dan tidak saling bergantung). Algoritma ini dipilih sebagai pembanding karena efisiensinya dalam memproses *dataset* besar dengan kecepatan komputasi yang tinggi pada teks ulasan maupun opini publik (Yasir Alghifari dkk., 2025).

Pencarian kelas sentimen yang paling optimal dilakukan dengan membandingkan nilai probabilitas tertinggi dari setiap kelas, menggunakan persamaan berikut:

$$\hat{c} = \arg\max_{c \in C} P\,(c) \prod_{i=1}^{n} P(w_i|c)$$

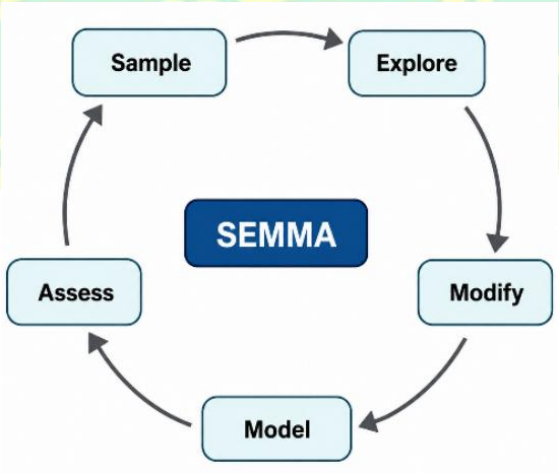
Keterangan:

1. $\hat{c}$: Kelas prediksi yang terpilih (kelas dengan nilai probabilitas tertinggi).
2. $\mathrm{P}(c)$: Probabilitas awal (prior) dari sebuah kelas sentimen.

3. $P(w_i|c)$: Probabilitas kemunculan sebuah kata $\{(w\}_i)$ jika diketahui kelas sentimennya adalah c.

2.7 Pendekatan SEMMA

Penelitian ini menggunakan kerangka kerja SEMMA (*Sample, Explore, Modify, Model, Assess*) sebagai metodologi sistematis. SEMMA menyediakan alur kerja yang terstandarisasi mulai dari pengambilan sampel data hingga tahap penilaian model. Penerapan SEMMA memastikan bahwa transisi dari data teks mentah menuju model prediksi dilakukan melalui tahapan yang terukur, terutama pada fase Modify (preprocessing) dan Model (implementasi algoritma), sehingga hasil komparasi antara SVM dan Naïve Bayes menjadi valid dan akurat. Penelitian oleh Saputra & Susanti (2025) juga menunjukkan bahwa SEMMA efektif diterapkan dalam analisis sentimen karena mampu mengintegrasikan proses crawling data, preprocessing, pemodelan menggunakan algoritma machine learning, hingga evaluasi model secara sistematis. Oleh karena itu, SEMMA dipilih sebagai kerangka kerja yang sesuai untuk mendukung proses penelitian ini



Gambar 2.1 Kerangka Metode SEMMA.

2.8 Evaluasi Peforma Model

Evaluasi model klasifikasi sentimen umumnya menggunakan *Confusion Matrix* dan metrik standar seperti:

1. *Accuracy*
2. *Precision*
3. *Recall*
4. *F1-score*

Menurut Leandro & Fianty (2025) serta Mao (2024), penggunaan kombinasi beberapa metrik evaluasi ini diperlukan untuk memberikan gambaran komprehensif mengenai kinerja algoritma dalam mendeteksi sentimen secara menyeluruh.

2.9 Penelitian Terkait

Beberapa penelitian yang relevan dengan topik perbandingan algoritma ini antara lain:

1. Yasir Alghifari (2025): Perbandingan algoritma svm dan naive bayes dalam analisis sentimen terhadap ulasan pengguna bukalapak
2. Rizki (2025): Perbandingan algoritma svm dan naive bayes dalam analisis sentimen terhadap ulasan pengguna tiktokshop.
3. Agustin & Nabil Nur Afrizal (2025): Implementasi kerangka kerja SEMMA untuk efisiensi pengolahan *data mining* secara iteratif.
4. Ibrahim dkk., (2025): Analisis sentimen public twitter terhadap kebijakan pemerintah menggunakan metode svm studi kasus ruu tni