

BAB II TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Sebagai bahan untuk memperkuat dan memudahkan penelitian tugas akhir ini, peneliti merangkum penelitian yang berhubungan dengan penelitian ini.

Penelitian yang dilakukan oleh Dedi Darwis, dkk dengan judul “Penerapan Algoritma Naïve Bayes untuk Analisis Sentimen *Review* Data Twitter BMKG Nasional” [16] melakukan pencarian komentar netral, komentar positif, dan komentar negatif terhadap data BMKG Nasional di Twitter menggunakan algoritma naïve bayes classier dengan hasil akurasi sebesar 68.97%. Penelitian yang dilakukan oleh M. I. Fikri, dkk dengan judul “Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter” [17]. Terdapat penelitian lainnya yang dijadikan referensi dalam penelitian ini seperti pada Tabel 2.1.

Pembaruan dalam penelitian ini adalah menggunakan data komentar dari YouTube dan membahas isu resesi tahun 2023 di Indonesia. Oleh karena itu peneliti membuat analisis sentimen terhadap isu resesi tahun 2023 di Indonesia dengan data komentar YouTube yang diambil pada tanggal 10 Oktober 2022 sampai 27 Oktober 2022 menggunakan algoritma Naïve Bayes.

Tabel 2.1 Penelitian Terkait

No	Judul Penelitian	Peneliti	Objek Penelitian	Metode Penelitian	Hasil
1	Penerapan Algoritma Naïve Bayes untuk Analisis Sentimen <i>Review</i> Data Twitter BMKG	Dedi Darwis, Nery Siskawati, Zenal Abidin	Melakukan pencarian netral, positif, dan negatif pada data Twitter BMKG Nasional	Algoritma Naïve Bayes Classifier	Berdasarkan hasil pengujian akurasi algoritma Naïve Bayes adalah sebesar 68.97%

No	Judul Penelitian	Peneliti	Objek Penelitian	Metode Penelitian	Hasil
	Nasional [16]				
2	Implementasi Sentimen Analisis Komentar <i>Channel Video</i> Pelayanan Pemerintah di YouTube menggunakan Algoritma Naïve Bayes [13]	Pramanan Yoga Saputra, Dian Hanifudin Subhi, Fahmi Zain Afif Winatama	Melakukan analisa sentiment terhadap kinerja pemerintah kepada masyarakat	Algoritma Naïve Bayes Classifier	Menurut hasil pengujian, algoritma Naïve Bayes Classifier menghasilkan akurasi sebesar 69.23% pada channel KemenPUPR dan 64.10% pada channel Kemenkeu
3	Analisis Sentimen Pembelajaran Daring pada Twitter di Masa Pandemi Covid-19 menggunakan Metode Naïve Bayes [20]	Samsir, Ambiyar, Unung Verawardina, Firman Edi, Ronal Watrianthos	Melakukan analisis opini publik tentang pendidikan online selama pandemi Covid-19 di Indonesia	Algoritma Naïve Bayes	Metode Naive Bayes menunjukkan 30% sentimen positif, 69% sentimen negatif, dan 1% sentimen netral dalam penelitian ini
4	Analisis Sentimen terhadap Opini Masyarakat Tentang Vaksin Covid-19 menggunakan Algoritma Naïve Bayes	Winda Yulita, Eko Dwi Nugroho, Muhammad Habib Algifari	Analisis opini masyarakat Indonesia tentang vaksinasi Covid-19	Algoritma Naïve Bayes Classifier	Hasil penelitian ini, opini masyarakat Indonesia—terutama di media Twitter—menunjukkan persentase positif

No	Judul Penelitian	Peneliti	Objek Penelitian	Metode Penelitian	Hasil
	Classifier [21]				sebesar 60.3%, dan algoritma Naïve Bayes Classifier menunjukkan akurasi sebesar 93%
5	Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter [17]	Mujaddid Izzul Fikri, Trifebi Shina Sabrila, Yufiz Azhar	Melakukan analisis sentimen opini masyarakat terhadap kampus Universitas Muhammadiyah Malah (UMM)	Algoritma Naïve Bayes dan Support Vector Machine (SVM)	Hasil pengujian menyatakan metode Naive Bayes memiliki hasil dengan nilai akurasi, ketepatan, recall, dan f1-score yang lebih baik dengan nilai 3.45%, 0.02, 0.04, dan 0.03 dibandingkan dengan metode SVM
6	Analisis Sentimen Pengguna YouTube terhadap Program Vaksin Covid-19 [12]	Ferian Fauzi Abdullah, Iqbal Rilo Pambudi	Mencari tahu opini masyarakat terhadap vaksin	Algoritma Support Vector Machine (SVM)	Hasil uji sentimen publik terhadap vaksin sebelum peristiwa vaksinasi yang dilakukan oleh Joko Widodo menunjukkan 41% netizen

No	Judul Penelitian	Peneliti	Objek Penelitian	Metode Penelitian	Hasil
					memiliki opini negatif tentang vaksin dan 17% memiliki opini positif. Namun, vaksinasi yang dilakukan oleh Joko Widodo dan setelahnya menunjukkan sikap negatif terhadap vaksinasi sebesar 24% dan sikap positif sebesar 36%
7	Analisis Sentimen Aplikasi Ruang Guru di Twitter menggunakan Algoritma Klasifikasi [11]	Angelina Puput Giovani, Ardiansyah, Tuti Haryanti, Lela Kurniawati, Windu Gata	Analisis sentimen dengan melakukan perbandingan algoritma klasifikasi terhadap aplikasi ruang guru	Algoritma Naïve Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbour (K-NN), dan <i>feature selection</i> dengan algoritma Particle Swarm Optimization (PSO)	Berdasarkan pengujian tersebut dapat diketahui bahwa algoritma SVM bekerja paling baik ketika pemilihan fungsi PSO digunakan dengan akurasi sebesar 78.55% jika dibandingkan algoritma NB dengan

No	Judul Penelitian	Peneliti	Objek Penelitian	Metode Penelitian	Hasil
					akurasi 67.32% dan K-NN dengan akurasi 77.21%
8	Analisis Sentimen Masyarakat Terhadap Kondisi Perekonomian di Indonesia pada Masa Pandemi 2020 [15]	Krisna Abhi Padhana, Mujiono Sadikin	Mencari tahu opini masyarakat khususnya di media Twitter mengenai ekonomi di Indonesia	Algoritma Naïve Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbour (K-NN)	Hasil penelitian menunjukkan bahwa algoritma SVM memiliki kinerja terbaik dengan akurasi 95,53%
9	Implementasi Naïve Bayes Classifier dan Confusion Matrix pada Analisis Sentimen Berbasis Teks pada Twitter [22]	Dwi Normawati, Surya Allit Prayogi	Mencari akurasi metode Naïve Bayes menggunakan data dari Twitter	Algoritma Naïve Bayes Classifier (NBC)	Hasil pengujian implementasi NBC dan performansi dengan Matriks menunjukkan akurasi sebesar 82%, ketepatan sebesar 93%, dan recall sebesar 52%
10	Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DRP) pada Twitter menggunakan Metode	Dianati Duei Putri, Gigih Forda Nama, Wahyu Eko Sulistiono	Mengetahui pendapat publik tentang kinerja Dewan Perwakilan Rakyat melalui media sosial Twitter	Algoritma Naïve Bayes Classifier (NBC)	Menurut penelitian yang dilakukan dengan algoritma Naive Bayes, skor akurasi sebesar 80%

No	Judul Penelitian	Peneliti	Objek Penelitian	Metode Penelitian	Hasil
	<i>Naïve Bayes</i> [23]				ditemukan, yang menunjukkan bahwa sistem dapat memprediksi 80% akurat dari 20% data pengujian

2.2 Dasar Teori

2.2.1 Isu Resesi 2023

Resesi adalah fenomena dimana PDB (Produk Domestik Bruto) riil berkurang selama dua triwulan berturut-turut, artinya kegiatan ekonomi seperti distribusi, investasi, konsumsi, produksi akan mengalami penurunan sehingga menimbulkan *domino effect* yang merugikan berbagai pihak, salah satunya adalah PHK (Pemutusan Hubungan Kerja) [1], [2]. Resesi ekonomi dapat didefinisikan sebagai suatu kondisi dimana PDB menurun atau pertumbuhan ekonomi riil yang rendah selama dua kuartal setiap tahun atau lebih [3]. Presiden Republik Indonesia Joko Widodo meminta agar masyarakat berhati-hati karena pada tahun 2023 diprediksi akan menjadi tahun yang gelap akibat krisis ekonomi dan energi akibat Covid-19 serta perang antara Rusia-Ukraina [6]. Menteri Keuangan Indonesia Sri Mulyani Indrawati menyatakan bahwa situasi ekonomi global sedang memburuk karena ancaman resesi ekonomi di semua negara, termasuk Indonesia. Namun, Dana Internasional memperkirakan pertumbuhan ekonomi Indonesia akan tetap 5% pada tahun 2022 dan 5% pada tahun 2023 [7].

2.2.2 Analisis Sentimen

Analisis sentiment atau *sentiment analysis* adalah metode yang digunakan untuk mengekstrak data opini, memahami, dan mengolah data tekstual secara otomatis untuk mengidentifikasi emosi dan perasaan yang terlibat dalam sebuah opini [24].

2.2.3 Data Mining

Data mining, juga dikenal sebagai penambangan data, adalah bidang yang mencakup banyak disiplin dan menggabungkan teknik pembelajaran mesin, basis data, statistik, pengenalan pola dan visualisasi untuk mengidentifikasi masalah dalam mengekstrak data dari basis data yang sangat besar. Dalam proses KDD (Knowledge Discovery in Databases), penambangan data termasuk pembersihan, integrasi, pemilihan, transformasi, penambangan, evaluasi pola, dan penyajian. [25]. *Data Mining* memiliki rangkaian proses tertentu yang harus dilakukan untuk memperoleh informasi baru, tahapan proses *data mining* adalah [26]:

1. *Data cleaning* atau Pembersihan data

Pembersihan data adalah proses untuk menghilangkan data yang tidak konsisten dan tidak relevan.

2. *Data integration* atau Integrasi data

Data yang dimasukkan ke dalam basis data baru dari berbagai basis data digabungkan melalui proses yang dikenal sebagai integrasi data.

3. *Data selection* atau Seleksi data

Karena informasi dalam basis data seringkali tidak dimanfaatkan secara maksimal, basis data hanya dicari informasi yang cocok untuk analisis atau yang cocok dengan data uji.

4. *Data transformation* atau Transformasi data

Data transformation dalam proses data mining artinya data diubah atau digabungkan menjadi format yang dapat diproses.

5. Proses *mining* atau penambangan

Saat menemukan informasi yang berharga dan tersembunyi dari data, proses mining adalah proses yang sangat penting.

2.2.4 Naïve Bayes

Dalam analisis sentimen, Naïve Bayes adalah metode klasifikasi dalam *text mining* yang berpotensi baik untuk klasifikasi dan komputasi data [20]. Pada algoritma Naïve Bayes, teorema Bayes digunakan untuk membuat rumus yang dapat digunakan untuk menghitung kemungkinan dari setiap klasifikasi yang mungkin dengan menggabungkan *prior probability* dan *conditional probability* [21]. Persamaan yang dapat digunakan untuk menghitung perbandingan antara *term* pada data uji dengan setiap kelas yang adalah [27]:

$$P(W_i|C) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|}$$

$$P(C|d) = P(C) \prod_{i=1}^n P(W_i|C)$$

C: kelas, d: dokumen, W_i : kata ke-i, $(W_i|C)$: jumlah kata w_i dalam C, Count (C): jumlah kata di kelas C, $|V|$: jumlah semua kata.

2.2.5 Evaluasi

Untuk melihat performa algoritma Naïve Bayes ada beberapa cara, dalam penelitian ini akan menggunakan Confusion Matrix untuk menghitung akurasi, presisi, recall, dan f1-score. Akurasi dihitung dengan membagi jumlah data sentiment benar dengan total data uji. Akurasi Naive Bayes dapat dihitung dengan menggunakan rumus berikut [23]:

$$\text{Akurasi} = \frac{\sum \text{data benar}}{n \text{ dokumen}} \times 100\%$$

Penghitungan tingkat presisi dilakukan dengan rumus sebagai berikut:

$$Presisi = \frac{\sum \text{data positif atau negatif}}{n \text{ dokumen positif atau negatif}} \times 100\%$$

Penghitungan tingkat *recall* dilakukan dengan rumus sebagai berikut:

$$Recall = \frac{\sum \text{dokumen relevan dan diambil}}{\sum \text{dataset positif atau negatif}} \times 100\%$$