e-ISSN : 2597-3673 (Online) , p-ISSN : 2579-5201 (Printed)
Vol.10 No.1 (June 2026)
Journal of Information System, Informatics and Computing
Website/URL: http://journal.stmikjayakarta.ac.id/index.php/jisicom
Email: jisicom@stmikjayakarta.ac.id , jisicom2017@gmail.com

ANALISIS KLASIFIKASI KUALITAS JAWABAN DEBAT POLITIK MENGGUNAKAN METODE *NAIVE BAYES* BERBASIS PRINSIP KERJA SAMA GRICE

Muhammad Tahir[1], Nurhafni[2], Bukran[3], Husain[4], Muhamad Wisnu Alfiansyah[5]

[1]D3-Sistem Informasi , Universitas Bumigora, Mataram
[2]Pendidikan Kepelatihan Olahraga, Universitas Bumigora, Mataram
[3]D3-Rekayasa Prangkat Lunak, Universitas Bumigora, Mataram
[4,5]Teknologi Informasi, Universitas Bumigora, Mataram

*Correspondent Author : muhammad_tahir@universitasbumigora.ac.id

Author Email: muhammad_tahir@universitasbumigora.ac.id [1], nurhafni@universitasbumigora.ac.id [2], bukran@universitasbumigora.ac.id [3], husain@universitasbumigora.ac.id [4], wisnu@universitasbumigora.ac.id [5]

In Indonesian

Abstrak: Penelitian ini bertujuan untuk menganalisis dan mengklasifikasikan kualitas jawaban debat politik menggunakan algoritma *Naive Bayes* berbasis Prinsip Kerja Sama Grice. Data diperoleh dari transkrip debat pasangan calon Bupati dan Wakil Bupati Bima Tahun 2024 yang dikembangkan menjadi 500 data teks dan diberi label ke dalam tiga kategori, yaitu Relevan, Menghindar, dan Tidak Relevan. Tahapan penelitian meliputi preprocessing teks, yaitu case folding, cleaning, stopword removal, dan stemming menggunakan Sastrawi, kemudian dilanjutkan dengan ekstraksi fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Proses klasifikasi dilakukan menggunakan algoritma *Multinomial Naive Bayes* dengan pembagian data 80% sebagai data latih dan 20% sebagai data uji. Hasil penelitian menunjukkan bahwa model menghasilkan 1.207 fitur TF-IDF dan mencapai akurasi sebesar 79,00%, precision 71,76%, recall 79,00%, serta F1-score 73,00%. Hasil menunjukkan bahwa metode *Naive Bayes* mampu mengidentifikasi pola kebahasaan yang berkaitan dengan kepatuhan maupun pelanggaran terhadap maksim percakapan Grice. Penelitian ini membuktikan bahwa pendekatan machine learning dapat digunakan sebagai alat bantu objektif untuk mengevaluasi kualitas komunikasi politik yang lebih efisien.

Kata kunci: Debat Politik; Naive Bayes; Prinsip Kerja Sama Grice; Klasifikasi Teks; Machine Learning;

In English

Abstract:. *This study aims to analyze and classify the quality of political debate responses using the Naive Bayes algorithm based on Grice's Cooperative Principle. The research data were obtained from the transcript of the 2024 Bima Regent and Vice Regent candidate debate, which was expanded into 500 text entries and labeled into three categories: Relevant, Evasive, and Irrelevant. The research process included text preprocessing, consisting of case folding, cleaning, stopword removal, and stemming using the Sastrawi library, followed by feature extraction using the Term Frequency–Inverse Document Frequency (TF-IDF) method. Classification was performed using the Multinomial Naive Bayes algorithm with an 80% training data and 20% testing data split. The results showed that the model generated 1,207 TF-IDF features and achieved an accuracy of 79.00%, precision of 71.76%, recall of 79.00%, and an F1-score of 73.00%. These findings indicate that the Naive Bayes method is capable of identifying linguistic patterns associated with both compliance and violations of Grice's conversational maxims. This study demonstrates that a machine learning approach can be utilized as an objective tool for evaluating the quality of political communication in a more efficient and measurable manner.*

Keywords: *Political Debate; Naive Bayes; Grice's Cooperative Principle; Text Classification; Machine Learning.*

e-ISSN : 2597-3673 (Online) , p-ISSN : 2579-5201 (Printed)
Vol.10 No.1 (June 2026)
Journal of Information System, Informatics and Computing
Website/URL: http://journal.stmikjayakarta.ac.id/index.php/jisicom
Email: jisicom@stmikjayakarta.ac.id , jisicom2017@gmail.com

Received: *2026-03-30****. Revised:*** *2026-06-26****. Accepted:*** *2026-06-30,* ***Issue Period:*** *Vol.10 No.1 (2026), Pp. 365-374*

I. PENDAHULUAN

Bahasa merupakan media utama atau alat komunikasi yang digunakan dalam interaksi antar individu untuk menyampaikan maksud dan tujuannya kepada orang lain[1]. Dalam konteks kenegaraan, Indonesia menganut sistem politik demokrasi di mana kedaulatan tertinggi berada di tangan rakyat. Salah satu perwujudan nyata dari kedaulatan ini adalah penyelenggaraan Pemilihan Kepala Daerah (Pilkada), yang merupakan momen vital untuk menentukan arah kebijakan dan pemerintahan daerah[2]. Pilkada bukan sekadar proses formal-prosedural, melainkan menjadi ajang pertarungan ide, gagasan, serta strategi komunikasi politik antar kandidat melalui media debat publik.

Debat politik merupakan instrumen demokrasi yang vital karena menjadi sarana bagi masyarakat untuk menilai kemampuan komunikasi, kualitas argumentasi, serta pemahaman kandidat terhadap isu-isu publik, Idealnya, seorang kandidat dituntut untuk memberikan jawaban yang jelas, relevan, informatif, dan sesuai dengan konteks pertanyaan yang diajukan oleh moderator[3].

Namun, dalam praktiknya, efektivitas komunikasi ini sering terhambat oleh berbagai permasalahan linguistik dan strategis[4]. Masalah utama yang ditemukan adalah banyaknya jawaban debat yang ambigu, tidak relevan, bahkan sengaja menghindari substansi pembahasan. Kandidat sering kali memberikan jawaban yang bersifat normatif atau terlalu umum tanpa menjelaskan solusi konkret terhadap permasalahan teknis yang ditanyakan.

Secara pragmatik, fenomena ini dikategorikan sebagai pelanggaran terhadap Prinsip Kerja Sama Grice, khususnya pada maksim relevansi, kuantitas, dan cara[5]. Misalnya, kandidat sering memberikan informasi berlebih yang tidak dibutuhkan untuk memperkuat argumen pribadi atau melakukan pengalihan isu guna menutupi ketidaktahuan substansi.

Selain itu, terdapat kendala dalam metode evaluasi kualitas komunikasi tersebut[6]. Selama ini, analisis terhadap kualitas jawaban debat politik berbasis prinsip pragmatik masih dominan dilakukan secara manual oleh peneliti bahasa. Proses manual ini memiliki kelemahan signifikan, yaitu membutuhkan waktu yang sangat lama dan memiliki tingkat subjektivitas yang tinggi dalam penafsirannya.

Perkembangan teknologi Natural Language Processing (NLP) memberikan peluang untuk mengatasi masalah tersebut melalui sistem otomatis[7]. Namun, klasifikasi otomatis terhadap teks debat formal memerlukan metode yang andal untuk menangani karakteristik bahasa yang kompleks.

Algoritma *Naive Bayes Classifier (NBC)* diusulkan karena memiliki performa yang baik, proses komputasi yang sederhana, cepat, serta sangat efektif untuk pengolahan data teks berukuran besar dengan asumsi independensi fitur.

II. METODE DAN MATERI

2.1. Prinsip Kerja Sama Grice (*Grice Cooperative Principle*)

Debat politik bukan sekadar forum adu argumen, melainkan instrumen vital dalam sistem demokrasi yang berfungsi sebagai media komunikasi publik utama bagi para kandidat untuk menyampaikan visi, misi, dan program strategis[7]. Melalui debat, masyarakat diberikan ruang terbuka untuk menilai secara langsung kemampuan komunikasi, kualitas argumentasi, serta kedalaman pemahaman seorang calon pemimpin terhadap isu-isu publik yang kompleks[8]. Setiap tuturan dan jawaban yang disampaikan kandidat dalam forum ini menjadi indikator krusial yang membentuk persepsi kolektif masyarakat mengenai integritas dan kualitas kepemimpinan.

Selain sebagai sarana edukasi politik, debat juga berperan sebagai ajang pertarungan ide dan strategi retorika yang digunakan kandidat untuk menarik simpati pemilih melalui pembangunan citra diri yang positif di mata publik. Kualitas komunikasi dalam sebuah interaksi sosial, termasuk debat politik, dapat diukur melalui

e-ISSN : 2597-3673 (Online) , p-ISSN : 2579-5201 (Printed)
Vol.10 No.1 (June 2026)
Journal of Information System, Informatics and Computing
Website/URL: http://journal.stmikjayakarta.ac.id/index.php/jisicom
Email: jisicom@stmikjayakarta.ac.id , jisicom2017@gmail.com

efektivitas pertukaran informasi berdasarkan Prinsip Kerja Sama Grice[9].Teori ini menekankan bahwa agar sebuah percakapan berjalan lancar, setiap penutur harus memberikan kontribusi yang sesuai dengan kebutuhan dan tujuan dari tahap percakapan tersebut berlangsung.

Prinsip ini dijabarkan ke dalam empat maksim utama[9]:

a. Maksim Kuantitas: Penutur diwajibkan memberikan informasi secukupnya sesuai kebutuhan mitra tutur, tanpa berlebihan atau memberikan data yang kurang memadai.
b. Maksim Kualitas: Mengharuskan penutur menyampaikan informasi yang benar, didukung oleh fakta-fakta yang meyakinkan, serta menghindari pernyataan yang diragukan kebenarannya.
c. Maksim Relevansi: Menetapkan bahwa setiap jawaban atau tuturan harus memiliki keterkaitan langsung dan selaras dengan konteks topik yang sedang dibicarakan
d. Maksim Cara: Menuntut penyampaian pesan yang jelas, runtut, tidak ambigu, dan tidak berbelit-belit guna menghindari kesalahpahaman.

Dalam realitas politik, kandidat sering kali melakukan pelanggaran maksimseperti memberikan jawaban normatif yang berbelit atau mengalihkan isu sebagai strategi retorika terencana untuk menghindari pertanyaan sulit atau menutupi ketidaktahuan terhadap substansi permasalahan.

2.2. Text Mining dan *Natural Language Processing (NLP)*

Text mining merupakan proses sistematis untuk menggali informasi bermakna dan pola-pola tersembunyi dari dokumen teks yang tidak terstruktur menggunakan teknik komputasi[10].

Dalam analisis debat politik yang melibatkan volume data besar, teknik NLP digunakan untuk mengotomatisasi pengolahan bahasa manusia melalui tahapan preprocessing yang ketat guna meningkatkan akurasi klasifikasi[11][12]. Proses ini meliputi Case Folding untuk menyeragamkan karakter teks, Cleaning untuk menghapus gangguan (noise) seperti URL dan tanda baca, Tokenizing untuk memecah kalimat menjadi unit kata, Stopword Removal untuk menghilangkan kata umum yang tidak informatif, serta Stemming untuk mengubah setiap kata berimbuhan menjadi kata dasar. Tahapan prapemrosesan ini sangat krusial karena data yang bersih dan terstruktur secara signifikan akan mendukung performa algoritma pembelajaran mesin dalam mengenali nuansa makna dalam jawaban kandidat.

2.3. Penelitian Terdahulu

Penelitian ini dibangun di atas landasan beberapa studi kunci yang telah mengeksplorasi integrasi pragmatik dan teknologi klasifikasi, Penelitian ini berupaya mengisi celah dengan menggabungkan standar penilaian linguistik Grice sebagai basis pelabelan dengan algoritma *Naive Bayes* untuk menciptakan sistem pemantauan kualitas jawaban debat yang objektif dan transparan di Kabupaten Bima. Adapaun penelitian terdahulu yang relean dengan penelitan ini dapat dilihat pada tabel 2 dibawah ini:

Tabel 2 : Penelitian Terdahulu

Aspek	Muhammad Tahir et al.(Debat Calon Bupati 2024)	Lola Enjelia et al. (Pilkada 2024)	Ridha Fauza et al. (Pilpres 2024)	Sofiana & Hermaliza (Pilpres 2019)
Objek Data	500 Transkrip Jawaban Debat Pilkada Bima	1.187 Tweet di Media Sosial X	Tweet di Media Sosial X	Tuturan Debat Pilpres (5 Sesi)
Algoritma	*Naive Bayes* (NBC) & TF-IDF	NBC vs K-NN & TF-IDF	NBC & *k-Fold Cross Validation*	Analisis Deskriptif Kualitatif
Output	Relevan, Tidak Relevan, Menghindar	Positif, Negatif, Netral	Positif dan Negatif	Pelanggaran 4 Maksim Grice
Akurasi	x	NBC (81,05%) unggul dari K-NN (75,26%)	Rata-rata 82,33% (Tertinggi 92,06% pada fold-10)	Relevansi (26), Kualitas (25), Cara (19), Kuantitas (14)

e-ISSN : 2597-3673 (Online) , p-ISSN : 2579-5201 (Printed)
Vol.10 No.1 (June 2026)
Journal of Information System, Informatics and Computing
Website/URL: http://journal.stmikjayakarta.ac.id/index.php/jisicom
Email: jisicom@stmikjayakarta.ac.id , jisicom2017@gmail.com

2.4. Metode

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis teks (text mining) dan pembelajaran mesin (machine learning) untuk mengklasifikasikan kualitas komunikasi politik. Fokus utama metodologi ini adalah mengintegrasikan teori Prinsip Kerja Sama Grice sebagai standar penilaian linguistik ke dalam algoritma *Naive Bayes Classifier* untuk pengolahan data otomatis.

1. **Sumber Data dan Karakteristik Dataset**

Data penelitian ini bersumber dari transkrip debat publik Pilkada Bima 2024 yang diperoleh melalui dokumentasi tayangan langsung dari chanel Youtube debat bima 2024. Dataset terdiri dari 500 entri data yang mencakup interaksi antara moderator dan pasangan calon. Data tersebut dikelompokkan ke dalam lima sub-tema strategis, yaitu:

Tabel 2 : Sub Tema Debat

Kategori	Pembahasan
SDA dan Lingkungan	Polusi kerusakan hutan dan krisis air
Tata Kelola Birokrasi	Profesionalisme ASN dan sistem rekrutmen
Ekonomi dan Infrastruktur	pembangunan pasar dan stabilitas ekonomi
Pendidikan	Penanganan anak putus sekolah dan beasiswa
Pertanian	kelangkaan pupuk dan kesejahteraan petani

Setiap entri data memuat atribut berupa sub-tema, pernyataan (pertanyaan moderator), jawaban kandidat dan tanggapan lawan.

2. **Pelabelan Data Berbasis Prinsip Grice**

Sebelum proses klasifikasi otomatis, terlebih dahulu dilakukan pelabelan data secara manual oleh ahli bahasa guna menjamin validitas dan reliabilitas dataset. Kriteria pelabelan didasarkan pada tingkat kepatuhan penutur terhadap empat maksim Grice (kuantitas, kualitas, relevansi, dan cara)[11]. Data dikategorikan ke dalam tiga label utama:

a. Relevan: Jawaban yang mematuhi maksim relevansi dan memberikan solusi konkret sesuai konteks
b. Tidak Relevan: Jawaban yang tidak sesuai dengan topik pertanyaan (pelanggaran maksim relevansi)
c. Menghindar: Jawaban yang bersifat normatif, ambigu, atau berbelit-belit (pelanggaran maksim cara dan kuantitas).

3. **Tahap Preprocessing Data**

Tahap ini bertujuan untuk membersihkan noise dan menstrukturkan data teks agar siap diproses oleh algoritma. Proses preprocessing meliputi[18]:

a. Case Folding: Mengubah seluruh karakter teks menjadi huruf kecil (lowercase).
b. Cleaning: Menghapus elemen yang tidak relevan seperti simbol, angka, dan tautan URL.
c. Tokenizing: Memecah kalimat menjadi unit kata atau token tunggal.
d. Stopword Removal: Menghapus kata-kata umum (seperti "dan", "yang", "di") yang tidak memiliki bobot informasi strategis.
e. Stemming: Mengubah kata berimbuhan menjadi kata dasar menggunakan library Sastrawi untuk menyeragamkan makna.

4. **Ekstraksi Fitur TF-IDF**

Setelah data bersih, diterapkan metode TF-IDF (Term Frequency-Inverse Document Frequency) untuk mengubah teks menjadi representasi numerik[13].

a. TF : Menghitung frekuensi kemunculan kata dalam satu jawaban debat.
b. IDF : Mengurangi bobot kata yang terlalu umum dan meningkatkan bobot kata. kunci yang spesifik (seperti "pupuk", "birokrasi", "beasiswa").

Bobot numerik ini berfungsi sebagai indikator penting bagi model dalam menentukan probabilitas kategori jawaban.

5. **Klasifikasi dengan *Naive Bayes Classifier***

Data kemudian dibagi menggunakan skema 80% untuk data latih (training) dan 20% untuk data

e-ISSN : 2597-3673 (Online) , p-ISSN : 2579-5201 (Printed)
Vol.10 No.1 (June 2026)
Journal of Information System, Informatics and Computing
Website/URL: http://journal.stmikjayakarta.ac.id/index.php/jisicom
Email: jisicom@stmikjayakarta.ac.id , jisicom2017@gmail.com

uji (testing). Algoritma yang digunakan adalah Multinomial Naive Bayes, yang bekerja berdasarkan prinsip probabilitas Teorema Bayes dengan asumsi independensi fitur. Model ini menghitung peluang posterior setiap jawaban untuk masuk ke kategori label tertentu (Relevan, Tidak Relevan, atau Menghindar) berdasarkan distribusi kata yang muncul dalam dataset pelatihan.

6. **Evaluasi Model**

Performa sistem diukur menggunakan Confusion Matrix untuk mengevaluasi akurasi prediksi dibandingkan dengan label asli dari ahli bahasa[14]. Metrik evaluasi yang digunakan meliputi:

a. Accuracy: Persentase total jawaban yang diklasifikasikan secara benar.
b. Precision: Tingkat ketepatan model dalam memprediksi kategori tertentu.
c. Recall: Kemampuan model dalam menemukan kembali seluruh data pada kategori tersebut.
d. F1-Score: Nilai rata-rata harmonis antara presisi dan recall untuk memberikan gambaran performa yang seimbang.

III. PEMBAHASA DAN HASIL

3.1. Hasil Pengolahan Data dan Preprocessing

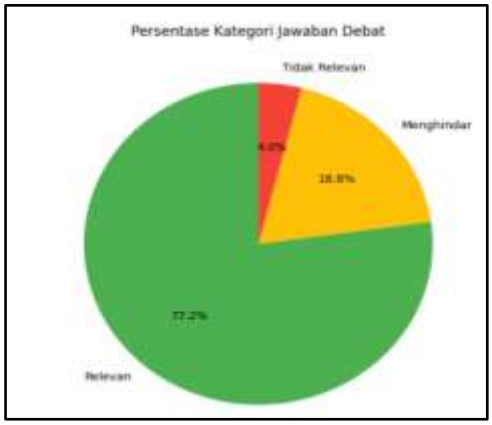
Penelitian ini menggunakan dataset transkrip debat pasangan calon Bupati dan Wakil Bupati Bima Tahun 2024 yang telah melalui proses pelabelan berdasarkan Prinsip Kerja Sama Grice. Dataset terdiri atas lima atribut utama yaitu nomor data (No), sub-tema, pernyataan, jawaban, dan label kualitas jawaban. Setelah dilakukan proses pembersihan data, diperoleh sebanyak 500 data yang siap digunakan dalam proses klasifikasi.

Tahapan preprocessing dilakukan untuk meningkatkan kualitas data teks sebelum dianalisis menggunakan algoritma machine learning. Proses ini meliputi case folding untuk menyeragamkan seluruh huruf menjadi huruf kecil, cleaning untuk menghilangkan karakter yang tidak diperlukan, stopword removal untuk menghapus kata-kata umum yang kurang informatif, serta stemming menggunakan pustaka Sastrawi untuk mengubah kata berimbuhan menjadi bentuk dasarnya.

Hasil preprocessing menunjukkan bahwa seluruh data dapat diproses dengan baik tanpa kehilangan informasi penting sehingga menghasilkan 500 data bersih yang siap digunakan pada tahap ekstraksi fitur dan klasifikasi.

3.2. Distribusi Kategori Kualitas Jawaban

Berdasarkan hasil pelabelan, distribusi data menunjukkan bahwa kategori Relevan merupakan kategori yang paling dominan dengan jumlah 386 data atau sekitar 77,2% dari total dataset. Kategori Menghindar berjumlah 94 data atau sekitar 18,8%, sedangkan kategori Tidak Relevan hanya berjumlah 20 data atau sekitar 4%. Hasil persentase dapat di lihat pada gambar di bawah ini;



Gambar 1 : Persentase Kategori Jawaban

e-ISSN : 2597-3673 (Online) , p-ISSN : 2579-5201 (Printed)
Vol.10 No.1 (June 2026)
Journal of Information System, Informatics and Computing
Website/URL: http://journal.stmikjayakarta.ac.id/index.php/jisicom
Email: jisicom@stmikjayakarta.ac.id , jisicom2017@gmail.com

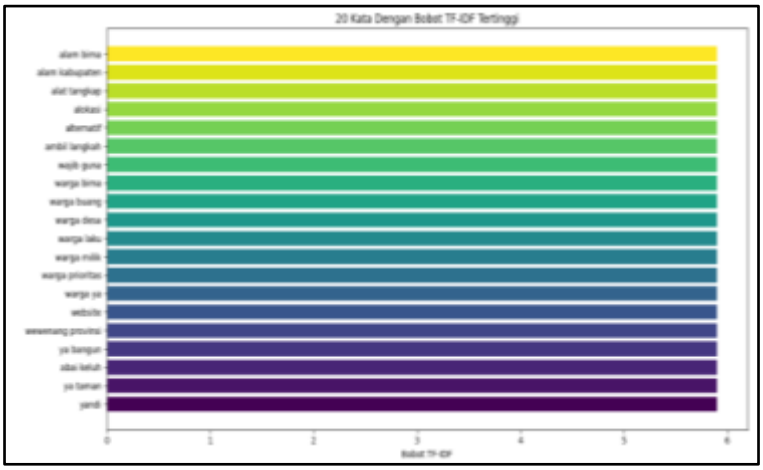
Distribusi ini menunjukkan bahwa sebagian besar jawaban kandidat masih berusaha menjawab substansi pertanyaan yang diberikan moderator maupun lawan debat. Namun demikian, masih ditemukan sejumlah jawaban yang cenderung menghindari substansi pertanyaan maupun tidak berkaitan langsung dengan topik yang dibahas.

Ketidakseimbangan jumlah data antar kategori menjadi salah satu faktor yang memengaruhi kemampuan model dalam mengenali seluruh kelas secara merata. Kategori Relevan yang mendominasi dataset menyebabkan model lebih banyak mempelajari pola bahasa pada kategori tersebut dibandingkan kategori lainnya.

3.3. Hasil Ekstraksi Fitur TF-IDF

Setelah preprocessing selesai dilakukan, tahapan berikutnya adalah ekstraksi fitur menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini digunakan untuk mengubah data teks menjadi representasi numerik sehingga dapat diproses oleh algoritma Naive Bayes.

Hasil ekstraksi fitur menghasilkan sebanyak 1.207 fitur unik yang berasal dari keseluruhan dokumen dalam dataset. Jumlah fitur tersebut menunjukkan bahwa jawaban debat memiliki keragaman kosakata yang cukup tinggi dan dapat menjadi indikator penting dalam proses klasifikasi kualitas jawaban.



Gambar 2 : Kata dengan TF-IDF Tertinggi

Melalui metode TF-IDF, kata-kata yang memiliki nilai informasi tinggi memperoleh bobot yang lebih besar dibandingkan kata-kata yang sering muncul pada seluruh dokumen. Dengan demikian, model dapat lebih mudah membedakan karakteristik bahasa pada kategori Relevan, Menghindar, dan Tidak Relevan.

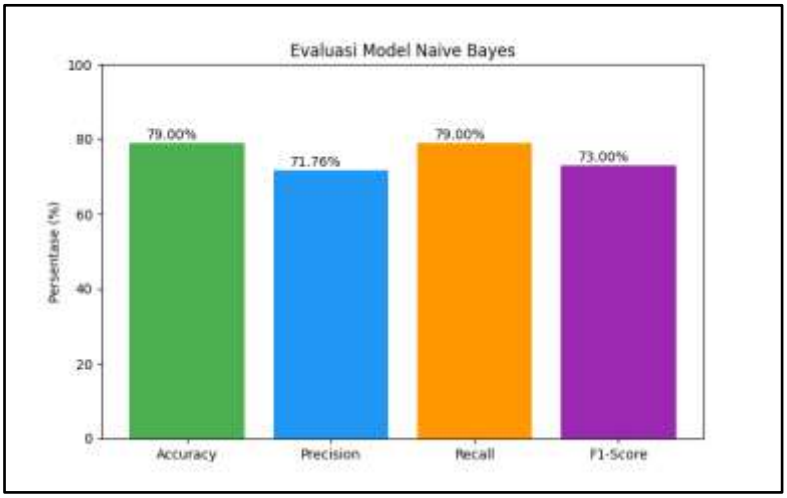
3.4. Hasil Klasifikasi Menggunakan Naive Bayes

Proses klasifikasi dilakukan menggunakan algoritma *Multinomial Naive Bayes* dengan skema pembagian data sebesar 80% untuk data pelatihan dan 20% untuk data pengujian. Dari total 500 data, sebanyak 400 data digunakan sebagai data latih dan 100 data digunakan sebagai data uji.

Setelah proses pelatihan selesai dilakukan, model digunakan untuk memprediksi kualitas jawaban pada data pengujian. Berdasarkan hasil evaluasi diperoleh nilai akurasi sebesar 79,00%. Hasil tersebut menunjukkan bahwa model mampu mengklasifikasikan 79 dari setiap 100 data uji dengan benar.

Selain akurasi, dilakukan pula pengukuran menggunakan metrik precision, recall, dan F1-score. Hasil pengujian menunjukkan nilai precision sebesar 71,76%, recall sebesar 79,00%, dan F1-score sebesar 73,00%.

e-ISSN : 2597-3673 (Online) , p-ISSN : 2579-5201 (Printed)
Vol.10 No.1 (June 2026)
Journal of Information System, Informatics and Computing
Website/URL: http://journal.stmikjayakarta.ac.id/index.php/jisicom
Email: jisicom@stmikjayakarta.ac.id , jisicom2017@gmail.com



Gambar 3 : Evaluasi Model Naive Bayes

Nilai recall yang lebih tinggi dibandingkan precision menunjukkan bahwa model cukup baik dalam menemukan data yang sesuai dengan kategori sebenarnya, meskipun masih terdapat sejumlah kesalahan klasifikasi terutama pada kategori yang jumlah datanya relatif sedikit.

3.5. *Analisis Classification Report*

Berdasarkan classification report, kategori Relevan memperoleh performa terbaik dibandingkan kategori lainnya. Kategori ini menghasilkan nilai precision sebesar 0,81, recall sebesar 0,99, dan F1-score sebesar 0,89. Hasil dari *Analisis Classification Report* dapat di lihat pada tabel di bawah ini ;

Tabel I. *Analisis Classification Report*

Classification Report	***Precisiaon***	***Recall***	***F1-Score***	***Support***
Menghindar	0.50	0.16	0.24	19
Relevan	0.81	0.99	0.89	77
Tidak Relevan	0.00	0.00	0.00	4
Accuracy			0.79	100
Macro avg	0.44	0.38	0.38	100
Weighted avg	0.72	0.79	0.73	100

Sumber : Hasil Procesing Python

Nilai recall sebesar 99% menunjukkan bahwa hampir seluruh jawaban yang benar-benar relevan berhasil dikenali oleh model. Hal ini menunjukkan bahwa pola bahasa pada kategori Relevan cukup konsisten sehingga mudah dipelajari oleh algoritma Naive Bayes.

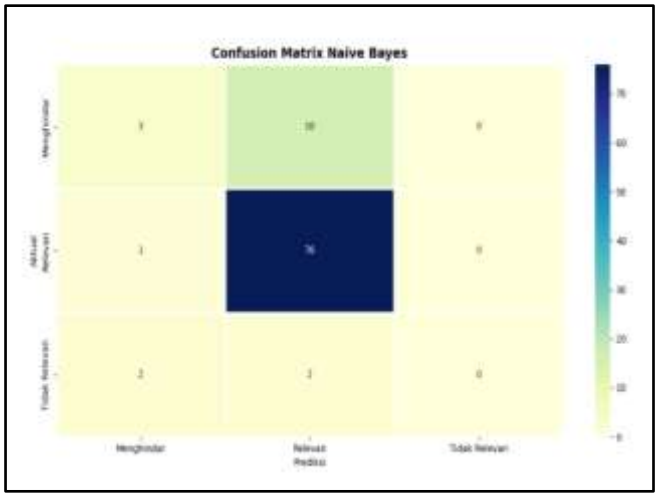
Pada kategori Menghindar diperoleh precision sebesar 0,50, recall sebesar 0,16, dan F1-score sebesar 0,24. Hasil ini menunjukkan bahwa model masih mengalami kesulitan dalam membedakan jawaban menghindar dengan jawaban relevan. Secara pragmatik, kondisi tersebut dapat dipahami karena jawaban menghindar sering kali masih memiliki hubungan dengan topik pembicaraan meskipun tidak menjawab inti pertanyaan secara langsung.

Sementara itu, kategori Tidak Relevan memperoleh nilai precision, recall, dan F1-score sebesar 0,00. Kondisi ini terjadi karena jumlah data kategori Tidak Relevan sangat sedikit dibandingkan kategori lainnya sehingga model tidak memiliki cukup contoh untuk mempelajari karakteristik bahasa pada kategori tersebut secara optimal.

3.6. *Analisis Confusion Matrix*

e-ISSN : 2597-3673 (Online) , p-ISSN : 2579-5201 (Printed)
Vol.10 No.1 (June 2026)
Journal of Information System, Informatics and Computing
Website/URL: http://journal.stmikjayakarta.ac.id/index.php/jisicom
Email: jisicom@stmikjayakarta.ac.id , jisicom2017@gmail.com

Berdasarkan hasil confusion matrix diperoleh bahwa dari 77 data yang sebenarnya termasuk kategori Relevan, sebanyak 76 data berhasil diprediksi dengan benar oleh model dan hanya 1 data yang salah diklasifikasikan sebagai Menghindar. Hasil ini menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam mengenali jawaban yang sesuai dengan substansi pertanyaan.



Gambar 4 : *Confusion Matrix*

Pada kategori Menghindar, dari 19 data yang diuji hanya 3 data yang berhasil diklasifikasikan dengan benar, sedangkan 16 data lainnya diprediksi sebagai Relevan. Kondisi ini menunjukkan adanya kemiripan pola bahasa antara jawaban relevan dan jawaban menghindar sehingga model mengalami kesulitan dalam membedakan kedua kategori tersebut.

Untuk kategori Tidak Relevan, dari 4 data yang diuji tidak ada satu pun yang berhasil dikenali dengan benar. Sebanyak 2 data diprediksi sebagai Menghindar dan 2 data lainnya diprediksi sebagai Relevan. Hasil ini mengindikasikan bahwa jumlah data kategori Tidak Relevan yang sangat terbatas menyebabkan model belum mampu mempelajari pola karakteristik kelas tersebut secara memadai.

.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma *Multinomial Naive Bayes* yang dikombinasikan dengan metode ekstraksi fitur TF-IDF mampu digunakan untuk mengklasifikasikan kualitas jawaban dalam debat politik berdasarkan Prinsip Kerja Sama Grice. Dari total 500 data yang telah melalui tahapan preprocessing, diperoleh 1.207 fitur unik yang digunakan sebagai representasi teks dalam proses klasifikasi.

Hasil pengujian menunjukkan bahwa model menghasilkan akurasi sebesar 79,00%, precision sebesar 71,76%, recall sebesar 79,00%, dan F1-score sebesar 73,00%. Nilai tersebut menunjukkan bahwa pendekatan machine learning dapat dimanfaatkan sebagai alat bantu yang cukup efektif untuk mengevaluasi kualitas jawaban debat politik secara otomatis dan objektif.

Analisis lebih lanjut menunjukkan bahwa kategori Relevan memiliki performa klasifikasi terbaik dengan nilai recall mencapai 99%, yang menandakan bahwa model sangat baik dalam mengenali jawaban yang sesuai dengan substansi pertanyaan. Namun, performa pada kategori Menghindar dan Tidak Relevan masih relatif rendah. Kondisi ini dipengaruhi oleh ketidakseimbangan distribusi data, di mana kategori Relevan mendominasi dataset, serta adanya kemiripan pola bahasa antara jawaban Menghindar dan Relevan.

Penelitian ini membuktikan bahwa integrasi teori pragmatik Grice dengan algoritma *Naive Bayes* dapat digunakan untuk mengidentifikasi kualitas komunikasi politik dalam debat publik. Meskipun demikian, peningkatan jumlah dan keseimbangan data pada kategori Menghindar dan Tidak Relevan diperlukan agar performa model dapat ditingkatkan pada penelitian selanjutnya. Selain itu, penggunaan metode klasifikasi yang

e-ISSN : 2597-3673 (Online) , p-ISSN : 2579-5201 (Printed)
Vol.10 No.1 (June 2026)
Journal of Information System, Informatics and Computing
Website/URL: http://journal.stmikjayakarta.ac.id/index.php/jisicom
Email: jisicom@stmikjayakarta.ac.id , jisicom2017@gmail.com

lebih kompleks atau teknik penanganan data tidak seimbang juga dapat menjadi alternatif untuk memperoleh hasil yang lebih optimal.

Saran penulis untuk pengembngan selanjutnya agar Menyeimbangkan distribusi kelas data, Ketidakseimbangan jumlah data pada kategori Relevan, Menghindar, dan Tidak Relevan memengaruhi performa model, terutama pada kategori yang memiliki jumlah data sedikit. Maka, diperlukan penambahan data pada kategori Menghindar dan Tidak Relevan agar model dapat mengenali karakteristik masing-masing kelas dengan lebih baik dan membandingkan dengan algoritma lain agar penelitian berikutnya dapat membandingkan performa *Naive Bayes* dengan algoritma klasifikasi lainnya seperti Support Vector Machine (SVM), Random Forest, Decision Tree, maupun metode berbasis Deep Learning untuk mengetahui metode yang paling efektif dalam mengklasifikasikan kualitas jawaban debat politik.

REFERENASI

[1] U. A. Siregar, N. Silvi, and W. Hasibuan, "Bahasa Sebagai Alat Komunikasi Dalam Kehidupan Manusia," vol. 7, no. 2, pp. 95–104, 2022.

[2] M. N. Ummah, G. Susanto, and G. C. W. P, "Tindak Tutur komisif dalam debat publik pertama calon bupati dan calon wakil bupati Mojokerto pada pilkada tahun 2024," vol. 8, pp. 297–312, 2025.

[3] A. Kurniawan, "Prinsip kerja sama dalam komunikasi politik," J. Pendidik. Bhs. Indones., vol. 12, no. 1, 2023.

[4] W. Zhang, "Violation of cooperative principle in political debates," J. Lang. Polit., vol. 21, no. 2, pp. 250–268, 2022.

[5] R. Mukhtar, S. Chodijah, and E. Rahayu, "Pelanggaran prinsip kerja sama pada debat capres Republik Indonesia 2024," Triangulasi J. Pendidik. Kebahasaan, vol. 5, no. 1, 2024, doi: 10.55215/triangulasi.v5i1.5.

[6] G. D. Barbulet and A. I. Ursa, "Exploring the cooperative principle in cross-cultural contexts: A corpus-based pragmatic study," Languages, vol. 11, no. 2, p. 29, 2026, doi: 10.3390/languages11020029.

[7] L. Enjelia, Y. Cahyana, and D. Wahiddin, "Comparison of K-Nearest Neighbors and Naive Bayes Classifier Algorithms in Sentiment Analysis of 2024 Election in Twitter (X)," vol. 9, no. 3, pp. 946–954, 2025.

[8] A. R. Politik, D. Publik, K. Calon, W. Bupati, and K. Ponorogo, "Analisis Retorika Politik dalam Debat Publik Kedua Calon Bupati dan Wakil Bupati Kabupaten Ponorogo Tahun 2024," vol. 7, pp. 2803–2812, 2025, doi: 10.47476/reslaj.v7i9.8713.

[9] L. Education, "Pelanggaran Prinsip Kerja Sama dalam Debat Pemilihan Umum Calon Presiden 2019," vol. 1, pp. 95–104, 2021.

[10] A. A. Abdirahman, A. O. Hashi, U. M. Dahir, and M. Abdirahman, "Enhancing Natural Language Processing in Somali Text Classification : A Comprehensive Framework for Stop Word Removal," vol. 71, no. 12, pp. 40–49, 2023.

[11] T. M. Rizqi Amelia Putri, "Pemenuhan Dan Penyimpangan Prinsip Kerja Sama Grice: Podcast Deddy Corbuzier-Kuliah Itu Gak Penting," vol. 20, pp. 1–11, 2024.

[12] A. A. Nafea et al., "A Brief Review on Preprocessing Text in Arabic Language Dataset : Techniques and Challenges," vol. 2024, pp. 46–53, 2024, doi: https://doi.org/10.58496/BJAI/2024/007 I.

[13] M. Das, S. Kamalanathan, and P. Alphonse, "A Comparative Study on TF-IDF feature Weighting Method and its Analysis using Unstructured Dataset," vol. 5571, pp. 0–2, 2021.

[14] L. Zhu and D. Luo, "A Novel Efficient and Effective Preprocessing Algorithm for Text Classification," pp. 1–14, 2023, doi: 10.4236/jcc.2023.113001.

[15] P. Febriyani and A. Rabi, "Implikatur Percakapan Dalam Debat Pasangan Calon Bupati Dan Wakil Bupati Sambas Tahun 2020," pp. 1–9, 2020.

e-ISSN : 2597-3673 (Online) , p-ISSN : 2579-5201 (Printed)
Vol.10 No.1 (June 2026)
Journal of Information System, Informatics and Computing
Website/URL: http://journal.stmikjayakarta.ac.id/index.php/jisicom
Email: jisicom@stmikjayakarta.ac.id , jisicom2017@gmail.com

[16] R. Sistem, "Algoritma Multinomial Naïve Bayes Untuk Klasifikasi Sentimen Pemerintah Terhadap Penanganan Covid-19 Menggunakan Data Twitter," vol. 5, no. 10, pp. 820–826, 2021.

[17] A. Serlina and A. Rahim, "Comparative Analysis of Naïve Bayes Algorithm Performance in English and Indonesian Text Sentiment Classification on Duolingo Application in Playstore," vol. 14, no. March, pp. 165–171, 2025, doi: 10.34148/teknika.v14i1.1207.

[18] R. F. Majbur, F. Yanuar, D. Devianto, D. Matematika, and U. Andalas, "Penerapan Metode Naïve Bayes Classifier Dalam Menganalisis Sentimen Pada Media Sosial X Terhadap Pilpres 2024 Di Indonesia," vol. 5, no. 3, pp. 2012–2025, 2024.