Jurnal TICOM: Technology of Information and Communication
Volume 14, Nomor 2, Januari 2026 Hal: 80-84 E-ISSN: 2962-7982

Penerapan Metode Naive Bayes Classifier pada Klasifikasi Berita Google Alert RSS FEEDS

Nofiyani

Fakultas Teknologi Informasi, Sistem Informasi, Universitas Budi Luhur, Jakarta, Indonesia
Jl. Raya Ciledug, Petukangan Utara, Kebayoran Lama, Jakarta Selatan 12260
E-mail: nofiyani@budiluhur.ac.id
(*: corresponding author)

Abstrak— **Di masa modern ini jumlah publikasi berita setiap hari semakin meningkat yang menyulitkan pengguna dalam menemukan berita yang relevan sesuai kebutuhan atau minat pengguna. Oleh karena itu, diperlukan sistem yang mampu mengelompokkan berita secara otomatis. Klasifikasi berita merupakan salah satu penerapan text mining. Proses klasifikasi ini memerlukan metode yang efektif. Salah satu metode yang sering digunakan adalah Naive Bayes Classifier, karena dapat bekerja sangat baik dan memiliki tingkat akurasi yg lebih baik dibanding model classifier lainnya. Dengan memanfaatkan data yang diperoleh dari Google Alerts RSS Feeds yang terdiri dari 80 data yang terbagi menjadi 64 data latih dan 16 data uji. Hasil evaluasi model Confusion matrix menunjukkan bahwa metode Naive Bayes Classifier (NBC) mengklasifikasikan secara benar 11 sampel dari total 16 sampel data uji dengan nilai akurasi sebesar 68,75%, nilai presisi 66,25% dan nilai recall 66,67%.**

Kata Kunci— **Berita, Klasifikasi, Google Alerts, Confusion Matrix, NBC**

Abstract— ***In this modern era, the number of news publications is increasing every day, making it difficult for users to find relevant news according to their needs or interests. Therefore, a system is needed that can group news automatically. News classification is one application of text mining. This classification process requires an effective method. One method that is often used is the Naive Bayes Classifier, because it can work very well and has a better accuracy rate than other classifier models. By utilizing data obtained from Google Alerts RSS Feeds consisting of 80 data points divided into 64 training data and 16 test data points. The results of the Confusion Matrix model evaluation show that the Naive Bayes Classifier (NBC) method correctly classifies 11 samples out of a total of 16 test data samples with an accuracy value of 68.75%, a precision value of 66.25% and a recall value of 66.67%.***

Keyword— ***Confusion Matrix, Classification, Google Alerts, News, NBC***

I. PENDAHULUAN

Di zaman modern ini, hampir di setiap aspek kehidupan memanfaatkan teknologi, begitu juga dengan membaca [1] yang menyebabkan jumlah publikasi berita setiap hari semakin meningkat. Berbagai platform seperti RSS (really simple syndication) feeds menyajikan informasi secara real-time salah satunya melalui layanan Google Alert.

Banyaknya berita menyulitkan pengguna dalam menemukan berita yang relevan sesuai kebutuhan atau minat pengguna, serta membuat proses pengelompokan berita secara manual menjadi tidak efisien, memerlukan waktu lama, serta berpotensi menimbulkan kesalahan klasifikasi akibat subjektivitas manusia. Oleh karena itu, dibutuhkan sebuah sistem yang dapat mengelompokkan berita secara otomatis sesuai dengan kategori yang telah ditetapkan sebelumnya.

Klasifikasi berita merupakan salah satu penerapan text mining yang bertujuan untuk mengelompokkan dokumen teks ke dalam kategori-kategori yang telah ditetapkan. Proses klasifikasi ini memerlukan metode yang efektif dalam merepresentasikan teks serta algoritma yang mampu menghasilkan tingkat akurasi yang baik. Naive Bayes Classifier merupakan salah satu teknik klasifikasi yang sering digunakan, karena dapat bekerja sangat baik dan menunjukkan tingkat akurasi yang lebih tinggi dibandingkan teknik klasifikasi lainnya [2].

Dengan menerapkan metode Naive Bayes Classifier dalam melakukan klasifikasi berita yang diperoleh dari Google Alerts RSS feeds, diharapkan klasifikasi berita dapat dilakukan secara otomatis serta menghasilkan akurasi yang optimal.

II. METODE PENELITIAN

Tahapan penelitian yang akan dilakukan dalam penyusunan penelitian ini adalah sebagai berikut:

A. Pengumpulan Data

Dalam penelitian ini Google Alerts digunakan sebagai sumber data dalam tahapan pengumpulan data. Google Alerts merupakan salah satu layanan Google yang bisa dimanfaatkan untuk memonitor konten yang ada di dunia maya [3]. Dengan menetapkan kata kunci sesuai fokus penelitian dan melakukan pengaturan Google Alerts pada sumber berita daring, bahasa Indonesia, serta wilayah Indonesia dalam periode waktu yang telah ditentukan. Berita yang diperoleh kemudian didokumentasikan.

B. *Preprocessing* Teks

Preprocessing merupakan tahapan yang penting dalam pipeline data science yang memiliki dampak besar pada hasil akhir analisis [4]. Sedangkan menurut [5] *preprocessing* dalam pengelompokan data adalah tahapan yang sangat penting untuk memastikan data yang digunakan berkualitas, sesuai kebutuhan dan sudah siap untuk proses analisis lebih lanjut. Dimana terdapat 7 tahapan yang akan dilakukan yaitu case folding, noise removal, konversi slangword, stopword removal, steamming dan tokenization.

1) *Case Folding*

Merupakan proses penyeragaman semua elemen data menjadi huruf kecil [6]. Sedangkan menurut [7] tahap case folding adalah mengubah teks menjadi huruf kecil atau huruf besar agar perbedaan antara huruf besar dan kecil dalam dokumen teks menjadi lebih kecil.

2) *Noise Removal*

Menghilangkan tanda baca, angka, simbol, dan alamat web [8]. Sedangkan menurut [9] noise removal merupakan proses menghilangkan (emojis, URLs, numbers).

3) Konversi *slangword*

Konversi Slang Word adalah mengubah kata tidak baku atau singkatan menjadi kata baku dalam komunikasi [10].

4) *Stopword Removal*

Stopword Removal merupakan bagian tahapan *preprocessing* yang menghilangkan kata-kata yang tidak dianggap penting atau tidak memiliki makna [11]. Sedangkan menurut [12] Stopword removal adalah mengeliminasi kata yang tidak memiliki kontribusi dalam analisis dan menyisakan kata bermakna atau dianggap penting dalam proses klasifikasi data.

5) *Stemming*

Stemming merupakan rangkaian langkah yang bertujuan untuk mendapatkan bentuk dasar dari setiap kata dengan menghilangkan imbuhan yang ada dalam kata, baik itu sisipan, awalan, maupun gabungan keduanya yang terdapat pada kata turunan [13].

6) *Tokenization*

Tokenization merupakan proses pemisahan teks yang biasa disebut dengan token [14]. Sedangkan menurut [15] tokenization adalah proses pemisahan kalimat menjadi kata-kata.

C. Klasifikasi dengan Naive Bayes Classifier

Naive Bayes Classifier adalah teknik yang banyak digunakan karena mudah, cepat serta memiliki struktur yang sederhana, serta efektivitasnya tinggi. Konsep Naive Bayes membentuk prediksi kemungkinan di waktu yang akan datang dengan pengalaman di waktu lampau [14]. Bentuk umum rumus dapat dinyatakan dalam persamaan 1 dimana P(j| adalah probabilitas term t kategori j, P(j) adalah probabilitas dokumen berkategori j, dan P(t) adalah probabilitas kemunculan term t.

$$\mathrm{P(j \mid t)} = \frac{\mathrm{P(j)}P(\mathrm{t \mid t})}{P(\mathrm{t})} \quad (1)$$

D. Evaluasi Model

Confusion matrix merupakan visualisasi untuk menilai efektifitas model klasifikasi. Confusion matrix terdiri dari informasi kelas aktual atau sebenarnya dan kelas prediksi. Confusion matrix disajikan pada Tabel 1 berikut [16].

TABEL I
CONFUSION MATRIX

		Actual Class	
		Class = Yes	*Class = No*
Class Predictad	*Class = Yes*	*TP (True Positive)*	*FP (False Positive)*
	Class = No	*FN (False Negative*	*TN (True Negative)*

TP: Kondisi dimana data prediksi dan aktual bernilai positif.
FP: Kondisi dimana data tidak sesuai dengan nilai aktual.
FN: Kondisi dimana data prediksi bernilai negatif sedangkan data aktual positif.
TN: Kondisi dimana data prediksi dan aktual bernilai negatif.

Untuk menilai kinerja dari data yang diperoleh, kita menggunakan akurasi, presisi, serta recall. Yang dinyatakan dalam persamaan 2, 3, 4 [16]:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

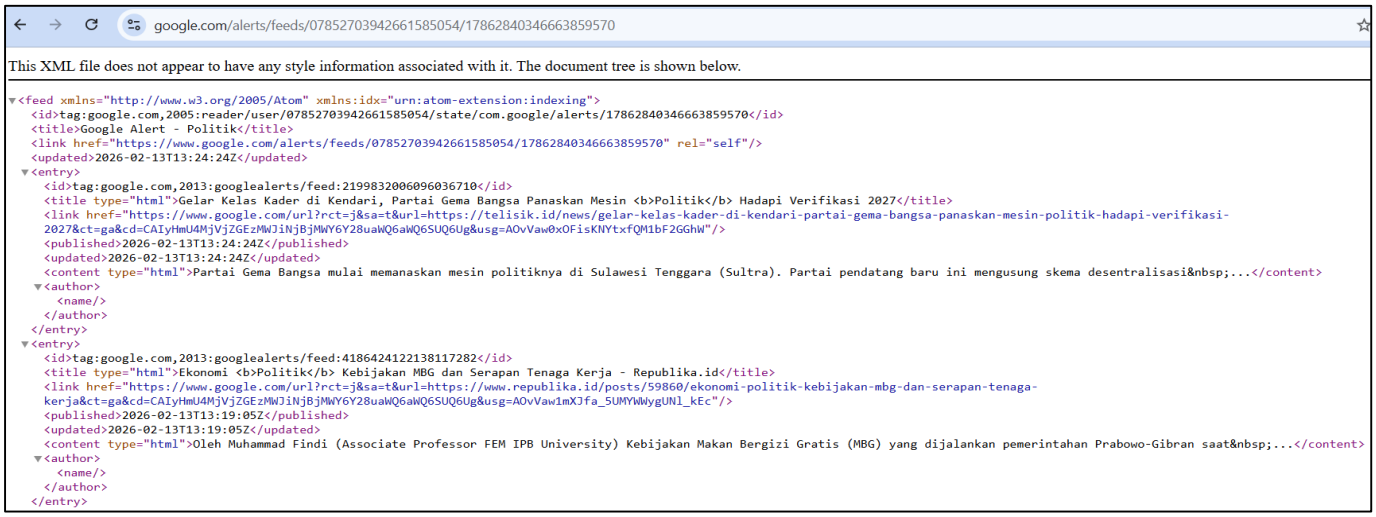
$$\text{Recall} = \frac{TP}{TP + FP} \quad (4)$$

III. HASIL DAN PEMBAHASAN

Dalam penelitian ini akan melakukan pengelompokan berita berdasarkan berita politik, ekonomi, hokum dan sosial. Ada beberapa tahapan yang akan dilakukan diantaranya:

A. Pengumpulan Data

Sumber data dalam penelitian ini diambil melalui RSS feeds Google Alerts terlihat pada Gambar 1, berdasarkan kata kunci berita yang sudah ditentukan. Beberapa sumber data yang akan digunakan adalah News, Blogs dan Web. Teks berita yang digunakan dalam penelitian diubah menjadi file XML, kemudian dikirim ke RSS feeds google alerts. URL feed RSS/XML yang didapatkan dimasukkan dalam aplikasi untuk memproses konten feed (Gambar 2).



Gambar 1. Tampilan File XML Google Alerts

Kembali Ke Form Input Link

No	ID	Title	Link	Publisher	Update	Content	Author
1	tag:google.com,2013:googlealerts/feed:2199832006096036710	Gelar Kelas Kader di Kendari, Partai Gema Bangsa Panaskan Mesin **Politik** Hadapi Verifikasi 2027		2026-02-13T13:24:24Z		Partai Gema Bangsa mulai memanaskan mesin politiknya di Sulawesi Tenggara (Sultra). Partai pendatang baru ini mengusung skema desentralisasi ...	
2	tag:google.com,2013:googlealerts/feed:4186424122138117282	Ekonomi **Politik** Kebijakan MBG dan Serapan Tenaga Kerja - Republika.id		2026-02-13T13:19:05Z		Oleh Muhammad Findi (Associate Professor FEM IPB University) Kebijakan Makan Bergizi Gratis (MBG) yang dijalankan pemerintahan Prabowo-Gibran saat ...	
3	tag:google.com,2013:googlealerts/feed:14589739398481127392	Dulu Wakil, Kini Gubernur! Andi Sudirman Naik Helikopter ke Tanah **Politik** Nurdin Abdullah		2026-02-13T12:55:37Z		GUBERNUR SULSEL - Gubernur Sulawesi Selatan, Andi Sudirman Sulaiman, mencuri perhatian saat tiba di Kabupaten Bantaeng menggunakan helikopter hitam ...	
4	tag:google.com,2013:googlealerts/feed:7222414766377465683	Anutin Resmi Gandeng Partai Pendukung Thaksin, Peta **Politik** Thailand Berubah		2026-02-13T12:40:16Z		Jakarta, aktual.com – Perdana Menteri (PM) Thailand Anutin Charnvirakul mencapai kesepakatan koalisi dengan rival utamanya, Jumat (13/2/2026), ...	
5	tag:google.com,2013:googlealerts/feed:18000935700392268699	Pemilu Perdana Bangladesh Pulihkan Stabilitas **Politik** Negara - Kompas.id		2026-02-13T12:30:18Z		Kemenangan besar diraih Partai Nasionalis Bangladesh atau BNP dalam Pemilihan Parlemen yang digelar Kamis (12/2/2026).	
6	tag:google.com,2013:googlealerts/feed:11731732368237642368	Tolak Hukum Jadi Alat **Politik**, Prabowo: Saya Berani Kasih Amnesti dan Abolisi - Periskop		2026-02-13T12:23:36Z		Prabowo Subianto menegaskan supremasi hukum bersih tanpa intervensi **politik**, memastikan keadilan dan kepastian hukum demi stabilitas bangsa.	
7	tag:google.com,2013:googlealerts/feed:6119856529453456918	Kemenko Polkam Pastikan Pembangunan Bendung Maliana 2 Kedepankan Keadilan dan ...		2026-02-13T12:04:05Z		SIARAN PERS NO. 48/SP/HM.01.02/POLKAM/2/2026. Polkam, NTT – Kementerian Koordinator Bidang **Politik** dan Keamanan (Kemenko Polkam) mengoordinasikan ...	
8	tag:google.com,2013:googlealerts/feed:8831572762391456025	Golkar Terbitkan 4 Fatsun **Politik** untuk Kadernya, Apa Saja? - Tempo.co		2026-02-13T12:00:33Z		Empat fatsun tersebut disampaikan Sekretaris Jenderal Partai Golkar Muhammad Sarmuji dalam agenda Tasyakuran dan Refleksi Kebangsaan merayakan Puncak ...	
9	tag:google.com,2013:googlealerts/feed:17592328171661449570	Hasto Kristiyanto: Spirit Kemanusiaan di Lokasi Bencana adalah Jati Diri **Politik** PDI Perjuangan		2026-02-13T11:58:40Z		Mengabdi di SumateraMisi Kemanusiaan Lintas Provinsi PDI Perjuangaan Apresiasi Puluhan Tenaga Medis dan Sopir Ambulans Baguna.	
10	tag:google.com,2013:googlealerts/feed:15755151510454605242	Golkar menegaskan "fatsun **politik**": Dilarang serang kebijakan Prabowo - ANTARA News Ntt		2026-02-13T11:54:29Z		Ketua Fraksi Partai Golkar DPR RI Muhammad Sarmuji menegaskan fraksinya kini mempunyai "fatsun **politik**" yang di antaranya anggota Fraksi Golkar ...	
11	tag:google.com,2013:googlealerts/feed:5195396679976573492	Butuh Upaya Kolektif untuk Tingkatkan Peran Perempuan di Bidang **Politik** - Metro TV		2026-02-13T11:47:02Z		Jakarta: Wakil Ketua MPR Lestari Moerdijat (Rerie) menilai butuh upaya kolektif setiap anak bangsa dalam upaya meningkatkan peran aktif perempuan ...	
12	tag:google.com,2013:googlealerts/feed:9797452229361612638	PKB Fokus Konsolidasi Dan Strategi **Politik** - WASPADA		2026-02-13T11:44:46Z		13 Februari 2026 MEDAN (Waspada.id) – Dewan Pengurus Wilayah (DPW) Partai Kebangkitan Bangsa (PKB) Sumatera Utara secara resmi dikukuhkan untuk ...	
13	tag:google.com,2013:googlealerts/feed:12749306419952263297	Prabowo Tegaskan Hukum Tak Boleh Jadi Alat **Politik**: Saya Berani Berikan Abolisi dan Amnesti		2026-02-13T11:30:57Z		Presiden Prabowo Subianto menegaskan komitmennya terhadap penegakan hukum yang adil dan tidak tebang pilih. Presiden Prabowo Subianto menegaskan ...	
14	tag:google.com,2013:googlealerts/feed:15311043398867748805	DEPARTEMEN **POLITIK** DAN ILMU PEMERINTAHAN FISIP UNDIP GELAR BEDAH BUKU ...		2026-02-13T11:29:03Z		Kegiatan ini merupakan bagian dari komitmen FISIP UNDIP dalam mendorong tradisi akademik yang produktif dan relevan dengan dinamika **politik** ...	
15	tag:google.com,2013:googlealerts/feed:13044266954357865496	Empat Fatsun **Politik** Golkar Jaga Prabowo-Gibran, Ini Isinya - RMOL		2026-02-13T11:18:17Z		Partai Golkar memasang rem internal terkait sikap mendukung pemerintahan Prabowo Subianto-Gibran Rakabuming Raka. Empat fatsun atau sopan santun ...	
16	tag:google.com,2013:googlealerts/feed:6595260446823229809	Butuh Upaya Kolektif untuk Tingkatkan Peran Perempuan di Bidang **Politik** - Fraksi Nasdem		2026-02-13T11:13:35Z		JAKARTA (13 Februari): Butuh upaya kolektif setiap anak bangsa untuk merealisasikan sejumlah kebijakan dan langkah afirmasi yang ada, ...	
17	tag:google.com,2013:googlealerts/feed:7229840128503899215	Bawaslu Bali Soroti Kendala SIPOL, Verifikasi Keterwakilan Perempuan Terhambat		2026-02-13T10:57:15Z		Denpasar, Bawaslu Bali — Pemutakhiran data partai **politik** melalui Sistem Informasi Partai **Politik** (SIPOL) kembali menjadi titik tekan dalam upaya ...	
18	tag:google.com,2013:googlealerts/feed:7547804386752480830	Prabowo Tegaskan Hukum Bukan Alat **Politik**, Komitmen Tegakkan Keadilan Tanpa Keragu-raguan		2026-02-13T10:51:04Z		Prabowo berkomitmen menjaga kepastian hukum tanpa menjadikannya alat **politik**, demi menjamin stabilitas nasional dan keadilan bagi seluruh rakyat ...	
19	tag:google.com,2013:googlealerts/feed:11620824277146430488	Tegaskan Hukum Tak Boleh Jadi Alat **Politik**, Prabowo: Saya Berani Berikan Abolisi dan Amnesti		2026-02-13T10:25:28Z		Presiden Prabowo Subianto dalam Indonesia Economic Outlook 2026. Foto: Youtube Setpres. Tegaskan Hukum Tak Boleh Jadi Alat **Politik**, Prabowo: Saya ...	
20	tag:google.com,2013:googlealerts/feed:10696266383864288406	MAX Aplikasi Pesan Made in Rusia yang Dikondisikan secara **Politik** untuk Menyingkirkan WhatsApp		2026-02-13T10:22:37Z		Perusahaan raksasa Meta mengatakan Rusia secara agresif berupaya memblokir penuh layanan aplikasi pesan WhatsApp dan menggantikannya dengan MAX.	

Gambar 2. Tampilan Program Proses Penyimpanan Konten Feed Kategori Berita Ekonomi

B. *Preprocessing* Teks

Tahapan selanjutnya setelah data pengumpulan data adalah tahap *preprocessing text* yang terdiri dari 6 tahapan yaitu *case folding*, *noise removal*, konversi *slangword*, *stopword removal*, *steaming* dan *tokenization*. Data mentah yang terdiri dari berbagai bentuk diubah menjadi data yang terstruktur, terlihat pada Gambar 3 berikut.

Kembali Ke Form Awal

ID	Content	Case Folding	Hapus Simbol	Filter Slang Word	Filter Stop Word	Stimming	Tokenisasi
tag:google.com,2013:googlealerts/feed:10086520920371114693	PARTAI Demokrasi Indonesia Perjuangan atau PDIP menyatakan 21 butir sikap **politik** usai menyelenggarakan rapat kerja nasional I yang dihelat sejak ...	partai demokrasi indonesia perjuangan atau pdip menyatakan 21 butir sikap **politik** usai menyelenggarakan rapat kerja nasional i yang dihelat sejak ...	partai demokrasi indonesia perjuangan atau pdip menyatakan butir sikap bpolitikb usai menyelenggarakan rapat kerja nasional i yang dihelat sejak	partai demokrasi indonesia perjuangan atau pdip menyatakan butir sikap bpolitikb usai menyelenggarakan rapat kerja nasional i yang dihelat sejak	partai demokrasi indonesia perjuangan pdip menyatakan butir sikap bpolitikb usai menyelenggarakan rapat kerja nasional i dihelat	partai demokrasi indonesia juang pdip nyata butir sikap bpolitikb usai selenggara rapat kerja nasional i helat	partai, demokrasi, indonesia, juang, pdip, nyata, butir, sikap, bpolitikb, usai, selenggara, rapat, kerja, nasional, i, helat
tag:google.com,2013:googlealerts/feed:10121020033571346238	Sebanyak 43 TAD menerima bantuan berupa paket sembako dan uang tunai yang disalurkan secara langsung sebagai wujud tanggung jawab **sosial** perusahaan ...	sebanyak 43 tad menerima bantuan berupa paket sembako dan uang tunai yang disalurkan secara langsung sebagai wujud tanggung jawab **sosial** perusahaan ...	sebanyak tad menerima bantuan berupa paket sembako dan uang tunai yang disalurkan secara langsung sebagai wujud tanggung jawab bsosialb perusahaan	sebanyak tad menerima bantuan berupa paket sembako dan uang tunai yang disalurkan secara langsung sebagai wujud tanggung jawab bsosialb perusahaan	tad menerima bantuan berupa paket sembako uang tunai disalurkan secara langsung wujud tanggung jawab bsosialb perusahaan	tad terima bantu upa paket sembako uang tunai salur cara langsung wujud tanggung jawab bsosialb usaha	tad, terima, bantu, upa, paket, sembako, uang, tunai, salur, cara, langsung, wujud, tanggung, jawab, bsosialb, usaha
tag:google.com,2013:googlealerts/feed:10427415723592735365	... **Sosial** Ketenagakerjaan, serta penandatanganan nota kesepakatan dengan BPJS Ketenagakerjaan. ... Program ini menyasar pekerja rentan, perangkat desa, ...	... **sosial** ketenagakerjaan, serta penandatanganan nota kesepakatan dengan bpjs ketenagakerjaan. ... program ini menyasar pekerja rentan, perangkat desa, ...	bsosialb ketenagakerjaan serta penandatanganan nota kesepakatan dengan bpjs ketenagakerjaan program ini menyasar pekerja rentan perangkat desa	bsosialb ketenagakerjaan serta penandatanganan nota kesepakatan dengan bpjs ketenagakerjaan program ini menyasar pekerja rentan perangkat desa	bsosialb ketenagakerjaan penandatanganan nota kesepakatan bpjs ketenagakerjaan program menyasar pekerja rentan perangkat desa	bsosialb ketenagakerjaan penandatanganan nota sepakat bpjs ketenagakerjaan program sasar kerja rentan perangkat desa	bsosialb, ketenagakerjaan, penandatanganan, nota, sepakat, bpjs, ketenagakerjaan, program, sasar, kerja, rentan, perangkat, desa
tag:google.com,2013:googlealerts/feed:10449426765335433752	OLEH Fahmi Salam Ahmad (Dosen Program Studi **Ekonomi** Pembangunan dan Sekretaris Departemen Ilmu **Ekonomi**, FEM IPB University) Sebagai seorang dosen ...	oleh fahmi salam ahmad (dosen program studi **ekonomi** pembangunan dan sekretaris departemen ilmu **ekonomi**, fem ipb university) sebagai seorang dosen ...	oleh fahmi salam ahmad dosen program studi bekonomib pembangunan dan sekretaris departemen ilmu bekonomib fem ipb university sebagai seorang dosen	oleh fahmi salam ahmad dosen program studi bekonomib pembangunan dan sekretaris departemen ilmu bekonomib fem ipb university sebagai seorang dosen	fahmi salam ahmad dosen program studi bekonomib pembangunan sekertaris departemen ilmu bekonomib fem ipb university dosen	fahmi salam ahmad dosen program studi bekonomib bangun sekertaris departemen ilmu bekonomib fem ipb university dosen	fahmi, salam, ahmad, dosen, program, studi, bekonomib, bangun, sekertaris, departemen, ilmu, bekonomib, fem, ipb, university, dosen
tag:google.com,2013:googlealerts/feed:11075717477084776324	Program KNMP Siap Perkuat Ketahanan Pangan hingga **Ekonomi** ... Selain mendukung MBG, KNMP berfungsi memperkuat **ekonomi** pesisir dengan menghubungkan hasil ...	program knmp siap perkuat ketahanan pangan hingga **ekonomi** ... selain mendukung mbg, knmp berfungsi memperkuat **ekonomi** pesisir dengan menghubungkan hasil ...	program knmp siap perkuat ketahanan pangan hingga bekonomib selain mendukung mbg knmp berfungsi memperkuat bekonomib pesisir dengan menghubungkan hasil	program knmp siap perkuat ketahanan pangan hingga bekonomib selain mendukung mbg knmp berfungsi memperkuat bekonomib pesisir dengan menghubungkan hasil	program knmp siap perkuat ketahanan pangan bekonomib mendukung mbg knmp berfungsi memperkuat bekonomib pesisir menghubungkan hasil	program knmp siap kuat tahan pangan bekonomib dukung mbg knmp fungsi kuat bekonomib pesisir hubung hasil	program, knmp, siap, kuat, tahan, pangan, bekonomib, dukung, mbg, knmp, fungsi, kuat, bekonomib, pesisir, hubung, hasil

Gambar 3. Tampilan Program Proses *Preprocessing Text*

C. Klasifikasi Data

Pendekatan analisis yang digunakan adalah Naive Bayes Classifier. Dimana akan melakukan pembagian data latih dan data uji. Pembagian data dalam penelitian ini adalah 80% sebagai data latih dan 20% data uji, yang terlihat pada Tabel II.

TABEL II
CLASSIFICATION DATA

Jumlah Data	Data Latih	Data Uji
80	64	16

D. Evaluasi Model

Tabulasi silang digunakan untuk menyajikan data aktual dan data prediksi model klasifikasi. Baris matriks menggambarkan kelas data aktual sedangkan kolom matriks menggambarkan kelas data hasil prediksi model, yang terlihat pada Tabel III berikut.

TABEL III
CONFUSION MATRIX NBC

Prediksi	Aktual			
	Politik	Hukum	Ekonomi	Sosial
Politik	4	0	0	1
Hukum	1	3	0	0
Ekonomi	1	1	3	0
Sosial	0	0	1	1

Berdasarkan tabel tersebut dapat dilihat hasil pengukuran kinerja model dengan melakukan perhitungan nilai *accuracy*, *precision* dan *recall* pada Gambar 4.

Menghitung Accuracy

==> Nilai Akurasi Yang Didapat : (11/16) x 100 = 68.75

Menghitung Precision

No.	Kategori	True Positive	False Positive	Precision
1	Politik	4	1	(4/ (4+1) = 0.8
2	Hukum	3	1	(3/ (3+1) = 0.75
3	Ekonomi	3	2	(3/ (3+2) = 0.6
4	Sosial	1	1	(1/ (1+1) = 0.5

==> Nilai Precision Yang Didapat : 66.25

Menghitung Recall

No.	Kategori	True Positive	False Negative	Recall
1	Politik	4	2	(4/ (4+2) = 0.66666666666667
2	Hukum	3	1	(3/ (3+1) = 0.75
3	Ekonomi	3	1	(3/ (3+1) = 0.75
4	Sosial	1	1	(1/ (1+1) = 0.5

==> Nilai Recall Yang Didapat : 66.666666666667

Gambar 4. Hasil Pengukuran Kinerja Model

Hasil penelitian menggunakan Naive Bayes Classifier terlihat berhasil mengklasifikasi secara benar 11 sampel dari total 16 sampel data uji dengan nilai akurasi sebesar 68,75%, nilai presisi 66,25% dan nilai recall 66,67%.

IV. PENUTUP

Kesimpulan berdasarkan hasil analisis untuk 80 data yang terbagi menjadi 64 data latih dan 16 data uji menunjukkan bahwa, frekuensi seberapa banyak model menghasilkan prediksi yang tepat dapat dilihat dari penghitungan akurasi sebesar 68,75%. Untuk hasil pengukuran proporsi hasil positif yang benar (TP) dari semua hasil diprediksi positif dapat dilihat dari perhitungan presisi sebesar 66,25%. Sedangkan hasil pengukuran proporsi hasil positif yang benar (TP) dari keseluruhan kasus sebenarnya positif terlihat dari perhitungan recall sebesar 66,67%.

UCAPAN TERIMA KASIH

Terima kasih penulis sampaikan pada editor dan reviewer untuk arahannya untuk perbaikan kualitas naskah. Terima kasih juga kepada keluarga atas dukungan dan doanya.

REFERENSI

[1] Repki, M. Fuad, and S. Samhati, “Manfaat Membaca Berita Bagi Siswa di SMK Swadhipa 2 Natar : Perspektif Aksiologi,” vol. 16, no. 2, pp. 67–78, 2024, doi: 10.30599/jti.v16i2.3298

[2] J. Sihombing, “Klasifikasi Data Antroprometri Individu Menggunakan Algoritma Naïve Bayes Classifier,” vol. 2, no. 1, pp. 1–10, 2021, doi: 10.37148/bios.v2i1.15

[3] J. Helianthusonfri, 10 Aplikasi Terbaik Google untuk Bisnis Anda, Elex Media Komputindo. Jakarta, 2020.

[4] I. M. Hamdani, *et al*., “Edukasi dan Pelatihan Data Science dan Data Preprocessing," *INTISARI: Jurnal Inovasi Pengabdian Masyarakat Edukasi*, vol. 2, no. 1, pp. 19–26, 2024, doi: 10.58227/intisari.v2i1.125.

[5] A. Agung, A. Daniswara, and I. K. D. Nuryana, “Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru,” vol. 5, no. 1, pp. 97–100, 2023, doi: 10.26740/jinacs.v5n01.p97-100.

[6] L. Rofiqi and M. Akbar, “Analisis Sentimen Terkait RUU Perampasan Aset dengan Support Vector Machine,” *JENIK: Jurnal Teknik Informatika*, vol. 4, no. 3, pp. 29-538, 2024, doi: 10.58794/jekin.v4i3.824.

[7] J. Jasmarizal, R. Rahmaddeni, J. Junadhi, and M. K. Anam, “Penerapan Metode Support Vector Machine untuk Analisis Sentimen Terhadap Produk Skincare,” vol. 13, no. 1, pp. 1438–1450, 2024, doi: 10.33022/ijcs.v13i1.3654

[8] A. F. Y. Khainur, et al., “Analisis Komparatif Efektivitas Pipeline Data Cleaning Berbasis Aturan dan Lemmatisasi untuk Klasifikasi Sentimen,” *Jurnal TIMES*, vol. 14, no. 2, pp. 141–149, 2025. [Online]. Available: https://ejournal.stmik-time.ac.id/index.php/jurnalTIMES/article/view/890/401

[9] H. Barus, I. N. Fajri, and Y. Pristyanto, “Sentiment Classification Analysis of Tokopedia Reviews Using TF-IDF, SMOTE, and Traditional Machine Learning Models,” *Journal of Applied Informatics and Computing*, vol. 9, no. 5, pp. 2552–2561, 2025, doi: 10.30871/jaic.v9i5.10524.

[10] F. Amrullah and A. Solichin, “Analisis Emosi Pada Live Chat Youtube ’ Mata Najwa : 3 Bacapres Bicara Gagasan Menggunakan Pendekatan Lexicon dan Algoritma Naive Bayes,” *Jurnal Ticom: Technology of Communication*, vol. 12, no. 3, pp. 121–128, 2024, 10.70309/ticom.v12i3.132.

[11] S. Dwi Prasetyo, S. ShofiahHilabi, and F. Nurapriani, “Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN,” *Jurnal KomtekInfo*, vol. 10, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.

[12] A. A. A. Pratamsyah and W. Widayat, “Analisis Sentimen Pengguna Twitter Terhadap Program Ibukotanusantara Menggunakan Metode Naïve Bayes,” Universitas Muhammadiyah Surakarta, 2025. [Online].

Available: https://eprints.ums.ac.id/132614/

[13] A. Sinaga, S. P. Nainggolan, "Analisis Perbandingan Akurasi Dan Waktu Proses Algoritma Stemming Arifin-Setiono dan Nazief-Adriani Pada Dokumen Teks Bahasa Indonesia," *Sebatik*, vol. 27, no. 1, pp. 63–69, 2023, doi: 10.46984/sebatik.v27i1.2072.

[14] D. F. Zhafira, B. Rahayudi, I. Indriati, "Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes dan Pembobotan TF-IDF Berdasarkan Komentar pada Youtube," *Jurnal Sistem Informasi Teknologi Informasi dan Edukasi Sistem Informasi*, vol. 2, no. 1, pp. 55–63, 2021, doi: 10.25126/justsi.v2i1.24.

[15] D. O. Sihombing, "Implementasi Natural Language Processing (NLP) dan Algoritma Cosine Similarity dalam Penilaian Ujian Esai Otomatis," *JSON: Jurnal Sistem Komputer dan Informatika*, vol. 4, pp. 396–406, 2022, doi: 10.30865/json.v4i2.5374.

[16] A. Nurul, Y. Salim, and H. Azis, "Analisis Performa Metode Gaussian Naïve Bayes Untuk Klasifikasi Citra Tulisan Tangan Karakter Arab," *Indonesia Journal of Data Science*, vol. 3, no. 3, pp. 115–121, 2022, doi: 10.56705/ijodas.v3i3.54.