

**SISTEM KLASIFIKASI BERITA KRIMINAL MENGGUNAKAN
METODE *NAÏVE BAYES CLASSIFIER***

SKRIPSI

**Oleh:
SAYYID MUHAMMAD UWAIS FATHONAH ASSAGAF
NIM. 19650052**



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2023**

**SISTEM KLASIFIKASI BERITA KRIMINAL MENGGUNAKAN
METODE *NAÏVE BAYES CLASSIFIER***

SKRIPSI

Diajukan kepada:
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana (UIN) Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan Dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh:
SAYYID MUHAMMAD UWAIS FATHONAH ASSEGAF
NIM. 19650052

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2023**

HALAMAN PERSETUJUAN

SISTEM KLASIFIKASI BERITA KRIMINAL MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

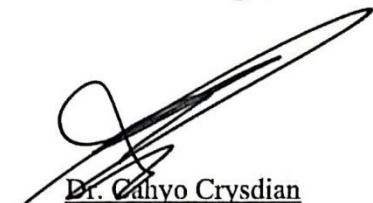
SKRIPSI

Oleh:

SAYYID MUHAMMAD UWAIS FATHONAH ASSEGAF
NIM. 19650052

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 30 November 2023

Pembimbing I,



Dr. Cahyo Crysdian
NIP. 19740424 200901 1 008

Pembimbing II,




Dr. Ririen Kusumawati, S.Si, M.Kom
NIP. 19720309 200501 2 002

Mengetahui,

Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi

Universitas Islam Negeri Maulana Malik Ibrahim Malang



Dr. Fachrudin Kurniawan, M.MT, IPM
NIP. 19771020 200912 1 001

HALAMAN PENGESAHAN

SISTEM KLASIFIKASI BERITA KRIMINAL MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

SKRIPSI

Oleh :

SAYYID MUHAMMAD UWAIS FATHONAH ASSEGAF
NIM. 19650052

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 21 Desember 2023

Susunan Dewan Penguji

Ketua Penguji : Dr. M. Amin Hariyadi, M.T
NIP. 19701182005011001

Anggota Penguji I : Fajar Rohman Hariri, M.Kom
NIP. 198905152018011001

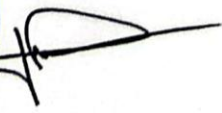
Anggota Penguji II : Dr. Cahyo Crysdian
NIP. 19740424 200901 1 008

Anggota Penguji III : Dr. Ririen Kusumawati, S.Si, M.Kom
NIP. 19720309 200501 2 002

()
()
()
()

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang




Dr. Fachrul Kurniawan, M.MT, IPM
NIP. 19771020 200912 1 001

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Sayyid Muhammad Uwais Fathonah Assegaf

NIM : 19650052

Fakultas / Jurusan : Sains dan Teknologi / Teknik Informatika

Judul Skripsi : Sistem Klasifikasi Berita Kriminal Menggunakan Metode
Naïve Bayes Classifier

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 21 Desember 2023

Yang membuat pernyataan,



Sayyid Muhammad Uwais F. Assegaf
NIM.19650052

HALAMAN MOTTO

"Keep moving forward, never look back."

"Selalu menyesuaikan diri dengan segala lingkungan."

"Pikiran positif, kehidupan lebih baik."

HALAMAN PERSEMBAHAN

**Puji syukur kehadiran Allah Subhanahu Wa Ta'ala, shalawat dan salam bagi
Rasul-Nya**

Skripsi ini dipersembahkan untuk kedua orang tua saya yaitu Suryani Nur dan Tajuddin Nur Khadir Assagaf sebagai bentuk tanggung jawab saya kepada orang tua. Perjuangan yang dilakukan oleh kedua orang tua saya semoga diberikan balasan yang setimpal oleh Allah SWT. Terima kasih juga saya sampaikan kepada orang-orang yang terlibat membantu secara langsung maupun tidak langsung dalam pembuatan skripsi ini.

KATA PENGANTAR

Assalamualaikum Warahmatullahi Wabarakatuh

Dengan menyebut nama Allah yang Maha Pengasih lagi Maha Penyayang, Segala puji dan syukur hanya bagi Allah Subhanahu Wa Ta'ala, Tuhan Semesta Alam, yang telah memberikan rahmat-Nya yang tiada terhingga serta petunjuk-Nya yang sempurna kepada hamba-Nya. Shalawat serta salam semoga tercurahkan kepada Nabi Muhammad Shallallahu 'Alaihi Wasallam, utusan Allah yang diutus sebagai rahmat bagi seluruh alam.

Dalam kesempatan yang berbahagia ini, dengan kerendahan hati dan penuh rasa syukur, saya dengan bangga mempersembahkan kata pengantar ini sebagai bagian dari penulisan skripsi berjudul “Sistem Klasifikasi Berita Kriminal Menggunakan Metode *Naïve Bayes Classifier*”. Skripsi ini telah saya susun untuk memenuhi syarat kelulusan dan mendapatkan gelar sarjana komputer (S.Kom) di Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Selanjutnya penulis mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu dan mendukung terselesaikannya skripsi ini. Ucapan terima kasih ini penulis sampaikan kepada :

1. Prof. Dr. H. M. Zainuddin, M.A., selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Dr. Sri Harini, M.Si, selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Fachrul Kurniawan, M.MT selaku Ketua Program Studi Teknik Informatika Universitas Maulana Malik Ibrahim Malang.

4. Dr. Cahyo Crysdiان dan Dr.Ririen Kusumawati,S.Si,M.Kom selaku dosen pembimbing yang dengan penuh kesabaran dan kebijaksanaan telah membimbing saya dalam menyelesaikan skripsi ini.
5. Ummi, Kakak, Adik-adik serta seluruh keluarga tercinta penulis yang turut memberi dukungan dan doa untuk penulis dalam menyelesaikan skripsi ini.
6. Dinda Natasya Putri Agung,S.Psi yang telah memberikan dukungan serta menemani penulis dalam menyelesaikan skripsi ini.
7. Segenap civitas akademika Jurusan Teknik Informatika, Terutama seluruh dosen, Terima Kasih atas ilmu dan bimbingan yang telah diberikan selama masa perkuliahan ini.
8. Teman-teman Jurusan Teknik Informatika Angkatan 2019 “Alien” yang sama-sama berjuang mengejar gelar S.Kom.
9. Penulis sendiri yang dapat menyelesaikan skripsi dan tanggung jawab penulis.
10. Semua pihak yang membantu dalam menyelesaikan skripsi ini, baik secara langsung maupun tidak langsung.

Akhir kata, penulis skripsi ini menyadari masih jauh dari ketidaksempurnaan. Oleh karena itu, penulis sangat mengharapkan saran dan kritik yang membangun untuk perbaikan selanjutnya. Semoga dalam penulisan skripsi ini banyak memberikan manfaat bagi berbagai pihak.

Malang, 21 Desember 2023

Penulis

DAFTAR ISI

HALAMAN PENGAJUAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
HALAMAN MOTTO	vi
HALAMAN PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
ABSTRAK	xiv
ABSTRACT	xv
خالصة.....	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Pernyataan Masalah.....	5
1.3 Tujuan Penelitian.....	5
1.4 Batasan Masalah.....	5
1.5 Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA	7
2.1 Klasifikasi Berita.....	7
2.2 <i>Natural Leanguage Processing (NLP)</i>	9
2.3 <i>Metode Naïve Bayes Classifier</i>	12
BAB III METODE PENELITIAN	15
3.1 Pengumpulan data	15
3.2 Desain Sistem	16
3.2.1 <i>Preprocessing</i>	17
3.2.2 <i>Pembobotan Kata Term Frequency</i>	23
3.2.3 <i>Naïve Bayes Classifier</i>	24
3.3 Implementasi sistem	28
BAB IV UJI COBA DAN PEMBAHASAN	32

4.1	Skenario Uji Coba	32
4.2	Hasil Uji Coba.....	34
4.3	Pembahasan	41
BAB V KESIMPULAN DAN SARAN		51
5.1	Kesimpulan.....	51
5.2	Saran.....	51
DAFTAR PUSTAKA		
LAMPIRAN		

DAFTAR GAMBAR

Gambar 1. 1 Jusmlah Kriminal dari BPS 2020	2
Gambar 3. 1 <i>Flowchart</i> Pengumpulan Data.....	16
Gambar 3. 2 Desain Sistem	16
Gambar 3. 3 <i>Flowchart Preprocessing</i>	17
Gambar 3. 4 Alur Proses <i>Case Folding</i>	18
Gambar 3. 5 Alur proses <i>Stopword removal</i>	20
Gambar 3. 6 Alur proses <i>stemming</i>	21
Gambar 3. 7 Alur Proses <i>Tokenizing</i>	22
Gambar 3. 8 Alur Proses TF	24
Gambar 3. 9 Alur Proses <i>Naïve Bayes Classifier</i>	25
Gambar 3. 10 Implementasi sistem.....	28
Gambar 3. 11 <i>Import Library Preprocessing</i>	29
Gambar 3. 13 Model <i>Multinomial Naïve Byes</i>	30
Gambar 3. 14 Hasil output	31
Gambar 4. 1 Grafik perbandingan hasil akurasi ketiga pengujian	41
Gambar 4. 2 Grafik perbandingan hasil presisi ketiga pengujian	42
Gambar 4. 3 Grafik perbandingan hasil recall ketiga pengujian.....	43
Gambar 4. 4 Grafik perbandingan hasil F-Measure ketiga pengujian	44

DAFTAR TABEL

Tabel 3. 1 Contoh <i>Case Folding</i>	19
Tabel 3. 2 Contoh <i>Stopword</i>	19
Tabel 3. 3 Contoh <i>Filtering (Stopword removal)</i>	20
Tabel 3. 4 Contoh hasil <i>Stemming</i>	22
Tabel 3. 5 Contoh <i>Tokenizing</i>	23
Tabel 4. 1 Skenario uji coba.....	32
Tabel 4. 2 Hasil klasifikasi pengujian 70:30.....	35
Tabel 4. 3 <i>Confusion Matrix</i> untuk pengujian 70:30	36
Tabel 4. 4 Hasil klasifikasi pengujian 80:20.....	37
Tabel 4. 5 <i>Confusion Matrix</i> untuk pengujian 80:20	39
Tabel 4. 6 Hasil klasifikasi pengujian 90:10.....	39
Tabel 4. 7 <i>Confusion Matrix</i> untuk pengujian 90:10	40

ABSTRAK

Assagaf, Sayyid M.U.F. 2023. **Sistem Klasifikasi Berita Kriminal Menggunakan Metode *Naïve Bayes Classifier***. Skripsi. Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Cahyo Crysdiyan (II) Dr. Ririen Kusumawati, S.Si, M.Kom.

Kata kunci: Berita Kriminal, Klasifikasi Berita, *Naïve Bayes Classifier*

Berita yang tersedia di media online lebih diminati oleh masyarakat karena lebih efisien dibandingkan dengan media cetak. Situs portal berita menjadi sumber informasi berita yang dapat menyajikan dan juga mengumpulkan berita-berita dari berbagai sumber untuk jadi landasan informasi kepada pembaca. Berita diunggah pada portal berita umumnya dikelompokkan dalam beberapa kategori, seperti berita politik, berita kriminal, atau berita lainnya. Namun, berita tersebut tidak ditempatkan di bawah sub kategori. Penelitian ini menggunakan salah satu metode klasifikasi teks yaitu metode *Naïve Bayes Classifier*. Tujuan dari penelitian ini untuk dapat mengukur seberapa akurat dalam mengklasifikasikan Berita kriminal, menggunakan metode ini dipilih untuk klasifikasi teks karena kemudahan implementasinya dan efisiensi komputasional yang dimilikinya. Data yang digunakan dalam metode *Naïve Bayes* merupakan data yang sudah melewati tahap *preprocessing*. Ada beberapa tahapan dalam *preprocessing* yakni *clean text* atau *case folding*, *tokenization*, *filtering (stopword removal)*, dan *Stemming*. Selanjutnya data yang telah melewati tahapan *preprocessing* akan dihitung tiap kemunculan kata pada tiap artikel berita menggunakan *Term-Frequency (TF)* dan nilai frekuensi inilah yang akan menjadi *input* pada metode *Naïve Bayes*. Pada pengujian pertama dengan rasio perbandingan 70:30 dengan akurasi 92,44%, presisi 81,11%, recall 81,11%, dan f-measure 81,11%. Hasil pengujian kedua dengan rasio perbandingan 80:20, dengan akurasi 91,33%, presisi 78,33%, recall 78,33%, dan f-measure 78,33%. dan hasil pengujian ketiga dengan rasio 90:10 menghasilkan akurasi 92%, presisi 80%, recall 80%, dan f-measure 80%. Hasil pengujian mengungkapkan bahwa kualitas data pelatihan memiliki dampak besar pada hasil akhir. Semakin besar jumlah data pelatihan yang digunakan, semakin baik hasilnya.

ABSTRACT

Assagaf, Sayyid M.U.F. 2023. **Criminal News Classification System Using Naïve Bayes Classifier Method.** Undergraduate Thesis. Department of Informatics Engineering Faculty of Science and Technology Maulana Malik Ibrahim State Islamic University Malang. Supervisor: (I) Dr. Cahyo Crysdiان (II) Dr.Ririen Kusumawati, S.Si, M.Kom.

Key words: *Criminal News, News Classification, Naïve Bayes Classifier*

News available in online media is more desirable to the public because it is more efficient than print media. News portal sites are a source of news information that can present and also collect news from various sources to become the basis of information to readers. News uploaded on news portals are generally grouped into several categories, such as political news, crime news, or other news. However, the news is not placed under sub-categories. This research uses one of the text classification methods, namely the Naïve Bayes Classifier method. The purpose of this research is to be able to measure how accurate it is in classifying criminal news, using this method was chosen for text classification because of its ease of implementation and the computational efficiency it has. The data used in the Naïve Bayes method is data that has passed the preprocessing stage. There are several stages in preprocessing, namely clean text or case folding, tokenization, filtering (stopword removal), and Stemming. Furthermore, the data that has passed the preprocessing stage will be calculated for each word occurrence in each news article using Term-Frequency (TF) and this frequency value will be the input to the Naïve Bayes method. In the first test with a ratio of 70:30 with 92.44% accuracy, 81.11% precision, 81.11% recall, and 81.11% f-measure. The second test results with a comparison ratio of 80:20, with an accuracy of 91.33%, precision of 78.33%, recall of 78.33%, and f-measure of 78.33%. and the third test results with a ratio of 90:10 resulted in 92% accuracy, 80% precision, 80% recall, and 80% f-measure. The test results reveal that the quality of the training data has a huge impact on the final results. The larger the amount of training data used, the better the results.

خالصة

السقاف ، سيد م.و.ف 2023 أطروحة. قسم Naïve Bayes يستخدم نظام تصنيف أخبار الجريمة طريقة تصنيف الهندسة المعلوماتية، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية، مالانج. مشرف (الأول) د. كاهيو كريسديان (الثاني) د. ريرين كوسوماواتي. بكالوريوس العلوم، ماجستير في علوم الحاسب الآلي

انكهد انشئسُخ: الأخبار الجنائية، تصنيف الأخبار، مصنف ساذج بايز

تعتبر الأخبار المتاحة في وسائل الإعلام عبر الإنترنت أكثر شعبية لدى الجمهور لأنها أكثر كفاءة مقارنة بالوسائط المطبوعة. تعد مواقع بوابات الأخبار مصادر للمعلومات الإخبارية التي يمكنها تقديم الأخبار وجمعها أيضًا من مصادر مختلفة لتصبح أساسًا للمعلومات للقراء. يتم عمومًا تجميع الأخبار التي يتم تحميلها على بوابات الأخبار في عدة فئات، مثل الأخبار السياسية أو أخبار الجريمة أو الأخبار الأخرى. ومع ذلك، لا يتم وضع الأخبار تحت أي فئة فرعية. يستخدم هذا البحث إحدى يهدف هذا البحث إلى قياس مدى دقتها في تصنيف أخبار *Naïve Bayes* طرق تصنيف النصوص وهي طريقة تصنيف الجرائم، وقد تم اختيار هذه الطريقة لتصنيف النصوص بسبب سهولة تنفيذها وكفاءتها الحاسوبية. البيانات المستخدمة في طريقة هي البيانات التي اجتازت مرحلة المعالجة المسبقة. هناك عدة مراحل في المعالجة المسبقة، وهي النص التنظيف *Naïve Bayes* أو طي الحالة، والتميز، والتصفية (إزالة كلمة التوقف)، والاستئصال. بعد ذلك، سيتم حساب البيانات التي اجتازت مرحلة وسيتم إدخال قيمة التردد هذه في طريقة (TF) المعالجة المسبقة لكل تكرار للكلمات في كل مقالة إخبارية باستخدام تردد المدى %، في الاختبار الأول، كانت النسبة 70:30 بدقة 92.44%، ودقة 81.11%، واستدعاء 81.11% *Naïve Bayes*. وكانت نتائج الاختبار الثاني بنسبة 80:20، وبدقة 91.33%، ودقة 78.33%، واستدعاء 81.11% f وقياس %، ونتائج الاختبار الثالث بنسبة 90:10 أتت بدقة 92% ودقة 80% واستدعاء 80% f وقياس 78.33%، أظهرت نتائج الاختبار أن جودة بيانات التدريب لها تأثير كبير على النتائج النهائية. كلما زادت كمية بيانات f 80% وقياس. التدريب المستخدمة، كانت النتائج أفضل

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi yang ditandai dengan adanya internet yang pesat mengakibatkan tingkat penyebaran informasi yang signifikan, dibuktikan oleh artikel yang dilansir dari (KOMINFO, 2022), dijelaskan bahwa pada tahun 2022 internet telah digunakan hingga 210 juta orang dengan persentasi 77.02 persen dari penduduk Indonesia, angka tersebut naik dibandingkan dengan tahun sebelumnya yaitu sekitar 73.7 persen (196.71 juta jiwa). Peranan penting dalam penyebaran informasi adalah media online seperti portal berita, ataupun media sosial. Media online saat ini dengan menyesuaikan tren membuat inovasi dalam menyampaikan berita berbasis data memunculkan suatu hal yang *up to date* ke dalam media *online*, sehingga berita tersebut dapat diakses dengan rentang waktu dan tempat tanpa adanya batasan (Al Fathan & Aminudin, 2021).

Berita yang tersedia dalam bentuk artikel di media *online* lebih diminati oleh masyarakat sebagai sarana untuk mengetahui tentang informasi yang telah terjadi dibanding dengan media koran, hal tersebut karena *efficiency*-nya dibanding dengan media cetak. Situs portal berita menjadi sumber informasi berita yang dapat menyajikan dan juga mengumpulkan berita-berita dari berbagai sumber untuk jadi landasan informasi kepada pembaca (Habib *et al.*, 2022). Berita yang ter-*publish* pada portal berita atau media *online* umumnya dikelompokkan dalam beberapa kategori, seperti berita politik, berita kriminal, atau berita lainnya. Namun, berita tersebut tidak ditempatkan di bawah sub kategori. Misalnya, jika pengguna mencari

kategori kriminal, berbagai berita kriminal ditampilkan tanpa indikasi jenis kejahatan yang tepat (Rhomadhona & Permadi, 2019).

Penelitian ini membahas mengenai klasifikasi berita kriminal, berdasarkan data yang penulis dapatkan dari Badan Pusat Statistik tahun 2022 dan dapat dilihat pada Gambar 1.1 menunjukkan jumlah total pelanggaran untuk tingkat Provinsi pada tahun 2021. Jumlah kejahatan terbesar dilaporkan di Kepolisian Daerah (Polda) Sumatera Utara (36.534), diikuti oleh Polda Metro Jaya (DKI Jakarta dan sekitarnya) dengan 29.103 kejahatan, dan Polda Jawa Timur (19.257). Dengan total 971 kejahatan, 1.000 kejahatan, 1.008 kejahatan, dan 1.500 kejahatan, Provinsi Kalimantan Utara, Maluku Utara, dan Sulawesi Barat merupakan tiga wilayah dengan jumlah kejahatan terendah.



Sumber: Biro Pengendalian Operasi, Mabes Polri

Gambar 1. 1 Jumlah Kriminal dari BPS 2020

Berdasarkan pemaparan di atas dibutuhkan klasifikasi yang menjadi metode untuk dapat menemukan sekumpulan model dalam mendeskripsikan dan membedakan kelas-kelas data sesuai dengan sub kategori pada setiap berita dengan kategori kriminal. Tujuan dari klasifikasi adalah untuk dapat melakukan prediksi kelas dari objek yang belum diketahui kelas serta karakteristik tipe datanya. *Preprocessing* merupakan proses atau tahapan yang dilakukan sebelum melakukan klasifikasi mulai dari *case holding*, setelah itu *tokenizing/parsing*, ketiga ada *filtering*, dan terakhir yaitu *stemming*. *Stemming* menjadi tahapan dalam *preprocessing* yang penting karena pada suatu kata akan dihilangkan imbuhan untuk menjadi kata dasar agar nantinya dapat memudahkan sistem dalam klasifikasi (Mustofa & Mahfudh, 2019).

Metode klasifikasi teks pun beragam mulai dari *Algoritme K-Nearest Neighbors* (KNN), *Naïve Bayes Classifier*, *Artificial Neural Network* (ANN), *Support Vector Machine* (SVM), *Random Forest*, *Decision Tree*, *Algoritme Logistic Regression*, *Multilayer Perceptron* (MLP), dan *Stochastic Gradient Descent* (SGD) (Sudianto *et al.*, 2022). Penelitian ini menggunakan salah satu metode klasifikasi teks yaitu metode *Naïve Bayes* dengan menggunakan *Term Frequency* (TF) sebagai pembobotan kata. Adapun penelitian terkait oleh Anshori (2020), melakukan penelitian untuk mengklasifikasikan kategori pada abstrak skripsi dengan metode *Naïve Bayes* dan juga metode *K-Nearest Neighbor* (KNN) sebagai pembanding. Hasil akhir dari penelitian ini menunjukkan bahwa Nave Bayes melebihi KNN dengan akurasi rata-rata 76,81%. Namun, jika hanya angka akurasi tertinggi yang digunakan, nilai KNN mengungguli nilai Nave Bayes. KNN

memberikan akurasi (tertinggi) sebesar 77.78% sedangkan Nave Bayes memberikan akurasi (tertinggi) sebesar 72.22%, selisih 5%. Selain dari nilai hasil akurasi rata-rata, pendekatan Nave Bayes melebihi pendekatan KNN dalam hal waktu pemrosesan hingga tiga kali lipat (0.31 detik berbanding 1.09 detik).

Penelitian ini diharapkan dapat mengamalkan perintah Allah dalam potongan Qur'an Surah Al-Baqarah ayat 197 *Allah Subhanahu wa ta'ala* berfirman:

الْحُجُّ أَشْهُرٌ مَّعْلُومَةٌ ۖ فَمَنْ فَرَضَ فِيهِنَّ الْحُجَّ فَلَا رَفَثَ وَلَا فُسُوقَ وَلَا جِدَالَ فِي الْحُجِّ ط وَمَا تَفَعَّلُوا مِنْ خَيْرٍ يَّعْلَمُهُ اللَّهُ ط وَتَزُودُوا فَإِنَّ خَيْرَ الزَّادِ التَّقْوَى ۗ وَاتَّقُوا يَا أُولِيَ الْأَلْبَابِ

“(Musim) haji itu (pada) bulan-bulan yang telah dimaklumi. Barangsiapa mengerjakan (ibadah) haji dalam (bulan-bulan) itu, maka janganlah dia berkata jorok (rafats), berbuat maksiat dan bertengkar dalam (melakukan ibadah) haji. Segala yang baik yang kamu kerjakan, Allah mengetahuinya. Bawalah bekal, karena sesungguhnya sebaik-baik bekal adalah takwa. Dan bertakwalah kepada-Ku wahai orang-orang yang mempunyai akal sehat!” (QS Al-Baqarah:197)

Ayat ini menekankan bahwa ibadah haji memiliki waktu yang pasti dan orang-orang yang ingin melaksanakannya harus mengikuti norma dan prosedur yang telah ditetapkan. Dalam konteks ini, klasifikasi dapat dipandang sebagai kepatuhan terhadap aturan dan prosedur tertentu. Sama halnya dengan bagaimana seorang Muslim mempersiapkan diri untuk menunaikan ibadah haji, sistem klasifikasi harus dipersiapkan dengan baik agar dapat menjalankan fungsinya secara efektif dan efisien. Sehingga penelitian ini berjudul “Sistem Klasifikasi Berita Kriminal Menggunakan Metode *Naïve Bayes Classifier*” sekiranya penelitian ini dapat menjadi kebijakan bagi masyarakat.

1.2 Pernyataan Masalah

1. Bagaimana cara melakukan proses klasifikasi berita kriminal pada data dari portal berita *online* menggunakan *Naïve Bayes Classifier*?
2. Sebarapa akurat hasil klasifikasi berita kriminal dengan menggunakan metode *Naïve Bayes Classifier*?

I.3 Tujuan Penelitian

1. Melakukan proses klasifikasi berita kriminal yang menggunakan data dari portal berita *online* dengan metode *Naïve Bayes Classifier*
2. Mengetahui keakuratan dari hasil klasifikasi berita kriminal menggunakan metode *Naïve Bayes Classifier*.

I.4 Batasan Masalah

- 1 *Dataset* untuk *data training* dari artikel berita bahasa Indonesia yang disimpan dalam format Excel secara manual
- 2 Berita yang dibutuhkan untuk pelatihan sebanyak 300 data yang diambil dari dua portal berita *online* yaitu Detik, dan Kompas dari bulan Juni 2022-Mei 2023.
- 3 Berita yang dikumpulkan hanya khusus berita dengan kategori kriminal, lalu memiliki 5 sub kategori yaitu pembunuhan, asusila, KKN(Korupsi, Kolusi, dan Nepotisme), pencurian, dan narkoba. Artikel berita hanya diklasifikasikan menjadi *single crime*.

1.5 Manfaat Penelitian

1. Bagi masyarakat, dapat mengetahui adanya pola permasalahan kriminalitas yang terjadi di masyarakat sehingga dapat lebih waspada.
2. Bagi kepolisian, dapat membantu dalam hal penangkapan pelaku kejahatan, pengurangan gangguan keamanan, dan evaluasi prosedur organisasi penegak hukum untuk didapatkan tindakan dan kebijakan yang tepat.
3. Bagi pemerintah, dapat dijadikan sebagai referensi dalam menentukan kebijakan dan pengambilan keputusan.
4. Hasil penelitian ini dapat dijadikan sebagai rujukan untuk penelitian-penelitian selanjutnya.

BAB II

TINJAUAN PUSTAKA

2.1 Klasifikasi Berita

Klasifikasi merupakan bahasa Perancis yang arti dalam bahasa Inggrisnya pun *classification*. Istilah ini merujuk pada tata cara pengaturan data secara teratur atau telah ditetapkan sebelumnya. Menurut Kamus Besar Bahasa Indonesia (KBBI), klasifikasi didefinisikan sebagai "proses pembentukan sistem pengelompokan berdasarkan ciri-ciri atau kriteria tertentu." Dalam hal ini, istilah tersebut mencakup langkah-langkah sistematis dalam menyusun data dan membentuk kelompok berdasarkan karakteristik atau standar yang telah ditetapkan sebelumnya. Di era digitalisasi, klasifikasi teks sangat penting karena ada perkembangan yang sangat besar dan eksplosif dalam pembuatan data teks. Karena ukuran data teks yang tersedia saat ini, terutama dengan maraknya media sosial, klasifikasi data menjadi penting untuk mengekstraksi pengetahuan yang dikandungnya. Bahan diklasifikasikan melalui proses klasifikasi ke dalam satu atau lebih kategori atau kelas bahan terkait yang telah ditentukan sebelumnya (Mutawalli *et al.*, 2019). Teks dikategorikan pada hakekatnya untuk mengklasifikasikan dokumen yang lazim digunakan pengkategorian teks contohnya seperti berita selama ini masih menggunakan tenaga manusia atau manual. Kategori yang banyak beserta waktu yang cepat akan sulit atau menjadi hambatan buat editor untuk mengkategorikan berbagai macam teks. Saat ini ada teknologi berupa *Machine Learning* dengan cara perumpamaan data latih, maka dapat melakukan klasifikasi dokumen secara otomatis (Juliana *et al.*, 2021).

Yudi dan Aditya (2020) melakukan penelitian terkait dalam melakukan klasifikasi teks berita memanfaatkan *Naive Bayes Classifier* dan algoritma yang disempurnakan *Confix Stripping Stemmer* untuk melakukan penelitian tentang klasifikasi berita. Peneliti dapat membedakan kata dasar dari prefiks dan imbuhan dengan menggunakan *Enhanced Confix Stripping Stemmer*. Saat ini, penulis menggunakan *Naive Bayes Classifier* untuk klasifikasi. Set data Informasi yang digunakan adalah 600 data dokumen berita di *website* www.jawapos.com. Informasi yang ada pada data ini dipartisi secara merata ke dalam beberapa kategori berita, termasuk teknologi, ekonomi, olahraga, dan kategori lainnya, yang akan dijadikan data latihan oleh penulis. Jumlah total data yang dianalisis mencapai 40, dengan tingkat akurasi masing-masing 90%, 90%, 100%, dan 100%, sehingga menghasilkan nilai total akurasi sekitar 95%.

Penelitian yang terkait lainnya oleh Dewi (2021) menggunakan metode *Multinomial Naïve Bayes* dalam pengujiannya menggunakan 10.500 *dataset* dan tujuh kategori pengujian pun diukur menggunakan *confusion matrix*. Hasil dari pengujian mendapatkan nilai skor yang baik tingkat *accuracy*, *precision*, *recall* dan *f1-score* memperoleh nilai 96%. Adapun penelitian lainnya dari Mustofa dan Mahfudh (2019), yang mengambil judul “Klasifikasi Berita *Hoax* Dengan Menggunakan Metode *Naïve Bayes*” data yang diambil diperoleh secara manual dari situs turnbackhoax.com dengan batasan berita *hoax* diambil dari bulan November 2018 sampai dengan Februari 2019. Total data sebanyak 300 yang dilabeli sebagai data latih dan tes. Proses klasifikasi ada beberapa tahapan mulai dari *preprocessing*, pembobotan kata dan dilakukan klasifikasi menggunakan

metode *Naïve Bayes*. Hasil menunjukkan dimana nilai *fold 6* memberikan nilai akurasi dengan hasil nilai keakuratan sebesar 85,28% dimana sebanyak 307 dokumen terklasifikasi relevan dan 53 dokumen tidak relevan atau *error rate* sebesar 14,72% dan juga hasil rata-rata berita *hoax* dan berita benar memiliki nilai *precision* 0,896 dan *recall* 0.853.

Pada penelitian ini akan menggunakan metode *Naïve Bayes Classifier* dengan TF-IDF sebagai pembobotan kata untuk klasifikasi teks. Adapun penelitian yang terkait hasil yang didapat dengan 900 *data train* lalu 100 *data test* dengan menerapkan metode *Naïve Bayes*, *TF-IDF* dan nilai *10-fold cross validation* menghasilkan nilai akurasi terbaik yaitu 97%. Berdasarkan hasil tersebut, metode *Naïve Bayes Classifier* dengan menggunakan pembobotan *TF-IDF* menunjukkan hasil yang cukup baik untuk alternatif klasifikasi (Zhafira *et al.*, 2021).

2.2 Natural Leanguage Processing (NLP)

NLP adalah cabang kecerdasan buatan yang berfokus pada pemrosesan bahasa alami. Bahasa alami adalah bahasa yang biasanya digunakan manusia untuk berkomunikasi satu sama lain. Bahasa yang dikenali oleh komputer harus diproses dan diinterpretasikan terlebih dahulu agar komputer dapat memahami maksud pengguna (Cucus *et al.*, 2019). Menurut penelitian oleh Sihombing (2022) yang mengimplementasikan NLP untuk digunakan melakukan proses data jawaban mahasiswa. Penelitian itu menggunakan NLP untuk mengolah teks dan melakukan *similarity* dengan algoritma *Cosine Similarity*. *Cosine Similarity* mengukur tingkat kemiripan antara dua dokumen menggunakan dua vektor dengan nilai mulai dari nol (0) hingga satu (1). Pemrosesan data jawaban dari siswa untuk tiga soal esai

untuk mendapatkan hasil yang diperlukan. Hasil perhitungan *Cosine Similarity* pada soal no. 1 menunjukkan bahwa siswa M3 memiliki respon dengan tingkat kesamaan 90.58%. Sedangkan untuk pertanyaan no. 2 siswa M1 memiliki respon dengan tingkat kemiripan 87.71%, dan untuk soal no. 3 siswa M1 memiliki jawaban dengan tingkat kemiripan 76.70%. Berdasarkan temuan tersebut, penerapan NLP dan algoritma *Cosine Similarity* dapat dimanfaatkan untuk menandai ujian esai secara otomatis.

Penelitian yang terkait lainnya oleh Prasetyo *et al.* (2021), mengimplementasikan bahwasanya dalam program teknologi informasi, NLP digunakan untuk membuat *chatbot*. Dengan adanya program teknologi informasi di Universitas Surabaya, Penulis mengembangkan suatu aplikasi chatbot yang bertujuan membantu pengguna dalam mendapatkan informasi relevan untuk penelitian mereka. Chatbot ini ditulis dalam bahasa Inggris dan dijalankan menggunakan perangkat lunak berbahasa Inggris, sehingga semua pertanyaan yang diterima harus dalam bahasa Inggris. Aplikasi chatbot yang dihasilkan menggunakan metode NLP untuk menganalisis pertanyaan yang diajukan oleh pengguna dan mengidentifikasi informasi kunci yang dibutuhkan. Sistem kemudian melakukan analisis pencarian informasi, dan jika informasi yang dibutuhkan tidak tersedia, chatbot akan memulai proses pencarian data dengan crawling untuk memperoleh informasi yang diinginkan oleh pengguna. Untuk validasi, penulis menggunakan metode validasi pengguna dan validasi silang. Validasi penulis metode silang menghasilkan akurasi sebesar 83%. Validasi pengguna penulis

menggunakan sepuluh titik data yang dikumpulkan dari pengguna untuk pengujian sistem, menghasilkan nilai akurasi 76% untuk para penulis.

Menurut penelitian oleh Megadiningrum *et al.* (2022), ada dua metode yang dibutuhkan dalam pembuatan sistem-nya. Metode NLP diimplementasikan untuk dapat memproses teks inputan atas keluhan anak. Kemudian ada metode *Forward Chaining* tersebut kemudian dimanfaatkan untuk mendapatkan perbaikan atau saran berdasarkan keluhan anak tersebut. Setelah 30 percobaan, metode pemrosesan bahasa alami mencapai akurasi 73% pada pencocokan kata kunci dan metode *Forward Chaining* mencapai akurasi 93% dalam menghasilkan kesimpulan sistem berdasarkan gejala trauma. Arsi *et al.* (2023) melakukan penelitian yang menggunakan NLP dan algoritma *machine learning* untuk menganalisis sentimen berdasarkan sentimen publik pada situs jejaring sosial twitter terkait jenis vaksin COVID-19 yang telah mendapatkan lisesnsi dari *World Health Organization* (WHO), tujuan dari penelitian tersebut untuk memastikan sentimen yang lazim di Twitter terhadap vaksin bersertifikat WHO Pfizer, Moderna, dan AstraZeneca berdasarkan NLP sebagai pembuat kebijakan dan sumber referensi untuk orang biasa. Menurut temuan penelitian, vaksin Pfizer jenis Moderna mendapat sikap paling positif, masing-masing sebesar 47.30% dan 46.20%. Sedangkan vaksinasi AstraZeneca memperoleh peringkat sentimen terendah (40.09%) dari ketiganya. Menurut temuan penelitian tersebut, konsumen lebih percaya pada jenis vaksin Pfizer dibandingkan jenis vaksinasi lainnya.

2.3 Metode *Naïve Bayes Classifier*

Algoritma *Naive Bayes Classifier* merupakan teknik klasifikasi sederhana yang disukai dan efektif berdasarkan teorema Bayes dan mengandaikan independensi di antara fitur yang digunakan untuk klasifikasi. Pendekatan ini, yang mengandalkan probabilitas, dianggap sebagai salah satu algoritma pembelajaran mesin yang paling mudah namun efisien (Indrayuni, 2019). Penelitian yang dilakukan Mandar et al. (2020) menggunakan metode *Naive Bayes classifier* untuk mengklasifikasi dokumen teks. Namun, sebelum diklasifikasikan, data berita yang berupa teks terlebih dahulu diproses dengan teknik *preprocessing* untuk menemukan indeks kata dalam berita berkualitas, seperti *case folding*, tokenisasi, *stopword*, dan *stemming*, dengan algoritma *stemming* yang digunakan adalah *porter stemmer*. Otomotif dan keuangan mendapatkan hasil yang lebih relevan dengan data pakar dari hasil pengujian terhadap 15 data berita yang dikategorikan oleh NBC dalam tiga kategori berita olahraga. Hasilnya, ditentukan bahwa penelitian ini memiliki akurasi sebesar 79% dalam mengklasifikasikan berita berdasarkan kategori atau rubrik.

Penelitian yang dilakukan oleh Nugroho et al. (2020) melakukan perbandingan dua metode antara *Naïve Bayes* dan *K-Nearest Neighbor* (K-NN) untuk mengetahui kinerja *Particle Swarm Optimization* (PSO). Penelitian tersebut mengklasifikasikan keluhan teks dengan lebih dari satu label pada satu waktu menggunakan NBC yang dioptimalkan menggunakan PSO. Data diperoleh dari portal sambat online dan dibagi menjadi 70% data latih dan 30% data uji untuk dikategorikan ke dalam tujuh label. Optimasi NBC dengan menggunakan PSO mencapai akurasi 87.44% dengan menggunakan *10-fold cross-validation*, lebih

tinggi dibandingkan dengan K-NN sebesar 75% dan NBC sebesar 64.38%, sehingga diharapkan dapat meningkatkan efektifitas layanan *e-government* bagi masyarakat.

Normawati & Prayogi (2021) melakukan penelitian yang menggunakan metode *Naive Bayes Classifier* (NBC) untuk memeriksa tingkat sentimen dari data yang di-*crawl* di Twitter. Sebagai studi kasus sederhana, hanya 8 *tweet* yang digunakan, dengan 5 *tweet* sebagai data pelatihan dan 3 *tweet* sebagai data uji. Data tersebut diproses terlebih dahulu sebelum dikategorikan menggunakan metode NBC dan performanya dihitung menggunakan teknik. Penelitian tersebut menggunakan *confusion matrix* untuk menghasilkan penjelasan sistematis mengenai proses dan hasil implementasi NBC dan penilaian kinerja, menghasilkan akurasi sebesar 82%, presisi 93%, dan recall 52%. Namun, hasil ini lebih mementingkan pembahasan setiap tahap dan metode secara lebih mendalam dibandingkan dengan angka-angka yang diperoleh. Kemudian, kumpulan data yang lebih besar dipelajari dengan menggunakan sistem aplikasi berbasis komputer.

Menurut penelitian Ruus *et al.* (2022) menggunakan metode Algoritma *Naive Bayes* untuk menghasilkan sistem analisis sentimen pada judul berita COVID-19 di media *online* dengan parameter positif dan negatif. Penelitian tersebut menggunakan pengukuran kinerja sistem untuk menilai kinerja aplikasi algoritma *Naive Bayes* dengan 50 data *training* dan 10 data *testing* berita. Sistem mengklasifikasikan 7 data *testing* dengan tepat dan salah mengklasifikasikan tiga data *testing*. Tiga buah data *testing* salah diprediksi. Akurasi metode klasifikasi adalah 0.7, presisi 0.6, dan *recall* 0.75. Hal ini menunjukkan bahwa metode

klasifikasi efektif untuk menilai dan mengkategorikan judul berita terkait COVID-19 di media *online*. Haris *et al.* (2022) melakukan penelitian yang menggunakan metode *Naïve Bayes* untuk dapat mengklasifikasikan warna kulit. Para penulis tersebut menggunakan koleksi foto wajah sebagai *dataset* untuk pelatihan dan pengujian *Naive Bayes*. *Python* adalah bahasa pemrograman yang digunakan untuk melakukan klasifikasi dengan *library* seperti *OpenCV* dan *scikit learn*. Akurasi maksimum yang didapatkan dari hasil uji coba menggunakan *Naive Bayes* adalah 75% dengan menggunakan *dataset* yang dimiliki penulis.

Penelitian yang serupa dilakukan Herfia Rhomadhona dan Jaka Permadi (2019), pada penelitian tersebut mengklasifikasikan berita kriminalitas berdasarkan sub kategori menggunakan metode *Naïve Bayes* (NBC) dokumen berupa artikel-artikel dari tiga sumber online di Kalimantan Selatan, seperti Banjarmasin.tribbunnews.com, antarakalsel.com, dan kalsel.prokal.co dengan presentase masing masing 34%, 33% dan 33% dokumen. Dengan total seluruh data 500 dokumen semua kategori. Penelitian tersebut menggunakan *K-Fold Cross Validation* untuk nilai diatur dari 3-10 menunjukkan hasil klasifikasi berita kriminal dengan nilai *precision* sebesar 98.53%, nilai *recall* sebesar 98.44% dan nilai *accuracy* sebesar 99.38.

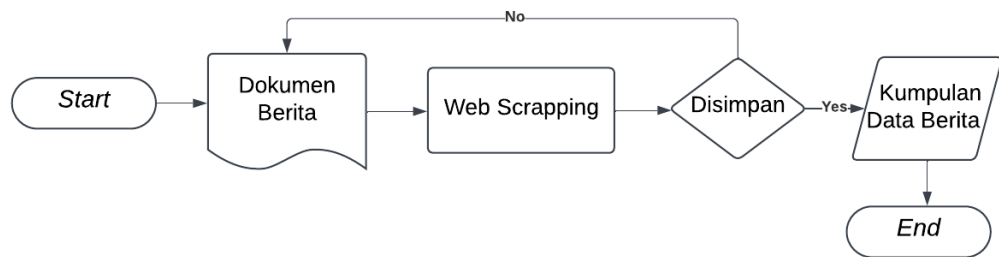
BAB III

METODE PENELITIAN

3.1 Pengumpulan data

Penelitian ini memanfaatkan dua tipe data, yakni data primer dan data sekunder. Data primer merupakan informasi yang dikumpulkan oleh peneliti secara langsung, sementara data sekunder adalah informasi yang diperoleh dari sumber luar. Sumber data sekunder termasuk portal berita online yang dapat diakses oleh masyarakat umum, seperti Kompas dan Detik. Setiap portal berita menyediakan berbagai jenis berita dan data yang dikumpulkan sesuai dengan batasan masalah yang ada pada latar belakang. Pengumpulan data dilakukan menggunakan *Tools* yaitu *Web Scraping* sehingga membentuk sebuah kumpulan data.

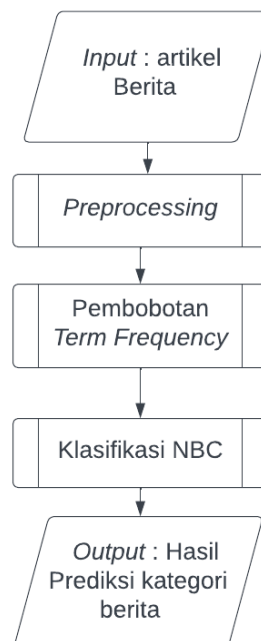
Web Scraping merupakan suatu proses menarik dokumen semi-terstruktur dari internet, biasanya dalam bentuk halaman *web* yang ditulis dalam bahasa markup seperti HTML atau XHTML, dan menganalisisnya untuk mengekstrak data spesifik dari halaman untuk digunakan dalam konteks lain (Turland, 2010). Data yang dikumpulkan menggunakan *Web Scraping* ada 300 berita yang terjadi di Indonesia dengan kategori kriminal dan akan diklasifikasikan berdasarkan sub-kategori yaitu pembunuhan, asusila, KKN, pencurian, dan narkoba seperti yang dapat dilihat pada Lampiran 2. Langkah langkah dalam melakukan pengumpulan data dapat dilihat pada gambar 3.1.



Gambar 3. 1 *Flowchart* Pengumpulan Data

3.2 Desain Sistem

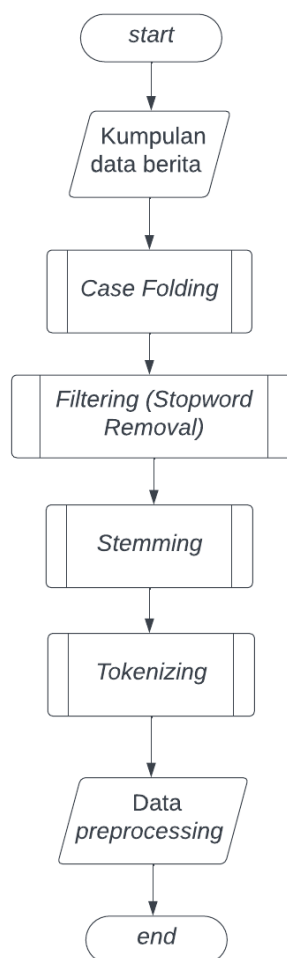
Desain sistem ini merupakan tahapan-tahapan dalam penelitian ini yang dapat digambarkan secara umum dari Gambar 3.2. Dalam membangun sistem untuk mengklasifikasikan Berita kriminal yang berbasis *web*. Tahapan proses terdiri dari kumpulan dokumen berita, tahap *preprocessing* yang didalamnya terdapat prosesnya sendiri (*case folding*, *tokenizing*, *Stopword*, *stemming*), pembobotan dengan *Term-Frequency* (TF), klasifikasi NBC, hasil klasifikasi dan selesai.



Gambar 3. 2 Desain Sistem

3.2.1 Preprocessing

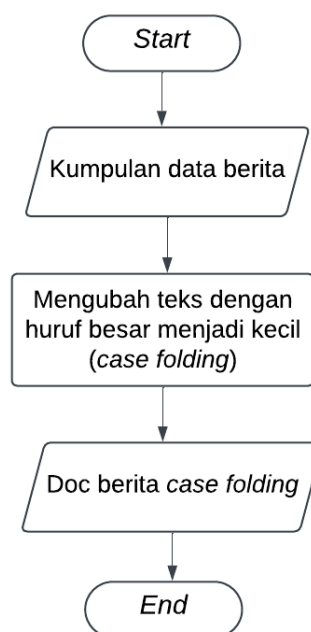
Pada tahap ini, data berita yang telah dikumpulkan sebelumnya akan diproses melewati empat proses yakni, *case folding*, *tokenizing*, *stopword*, dan *stemming*, yang tiap tahapannya dapat dilihat melalui Gambar 3.3 merupakan tahapan dari *preprocessing*. Berikut ini juga merupakan penjelasan mengenai *case folding*, *stopword*, *stemming*, dan *tokenizing*.



Gambar 3. 3 Flowchart Preprocessing

A. *Case Folding*

Case folding merupakan proses untuk mengubah semua huruf dalam teks dokumen menjadi huruf kecil adapun membersihkan data yang dilakukan untuk mengeliminasi tanda baca, simbol, angka, *emoticon* pada keseluruhan dokumen (Putra & Nugroho, 2021). Karena *string* bersifat *case sensitive*, maka Jika dibandingkan dengan istilah yang serupa namun tidak ditempatkan di awal frasa dan tidak menggunakan huruf besar, *string* tersebut akan diinterpretasikan dengan cara yang berbeda oleh sistem. Oleh karena itu, diperlukan *case folding* agar kata-kata yang semula memiliki struktur karakter yang berbeda dapat dianggap setara kembali, karena telah diubah menjadi bentuk yang serupa. Pada sistem proses *case folding* dibantu dengan menggunakan *library Natural Language Toolkit* (NLTK). Alur proses *case folding* dapat dilihat pada Gambar 3.4. Contoh input maupun hasil *output* dapat dilihat juga pada Tabel 3.1.



Gambar 3. 4 Alur Proses *Case Folding*

Tabel 3. 1 Contoh *Case Folding*

Contoh kalimat sebelum <i>case folding</i>	“Panglima TNI Laksamana Yudho Margono dan Kapolri Jenderal Listyo Sigit Prabowo meresmikan Markas Polda Papua yang terletak di Kampung Koya Koso”
Contoh Kalimat sesudah proses <i>case folding</i>	“panglima tni laksamana yudho margono dan kapolri jenderal listyo sigit prabowo meresmikan markas polda papua yang terletak di kampung koya koso”

B. *Filtering (Stopword)*

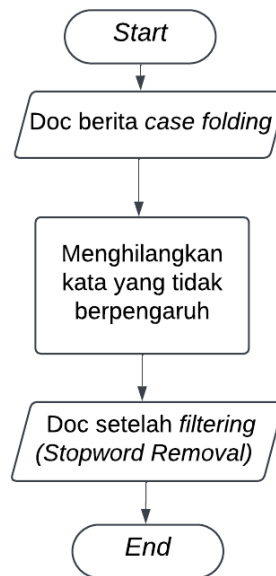
Salah satu prosedur *preprocessing* teks mining, memfilter *stopword removal*, termasuk menghapus kata-kata penghubung yang sering digunakan yang tidak memiliki signifikansi tertentu karena dapat memberikan pengaruh yang tidak baik dalam proses *text mining*. Tujuan utama dari *stopword removal* adalah untuk mempersingkat teks dan berkonsentrasi pada istilah yang lebih informatif dan membantu dalam pemahaman. Meskipun tiap *term* penghubung sering muncul dalam teks, *terms* tersebut biasanya tidak menambah banyak representasi makna atau fitur teks (Alita & Isnain, 2020). Seperti pada proses *case folding* untuk *filtering* dibantu menggunakan *library* NLTK. *list* kata yang ada menggunakan *library* NLTK dapat dilihat pada Lampiran 1. Pada Tabel 3.2 akan menampilkan beberapa contoh dari *stopword*.

Tabel 3. 2 Contoh *Stopword*

contoh <i>Stopword</i>	Yang	Dan	di
	dia	tak	lagi

Pada tahap ini, sistem akan mengambil kata-kata yang tidak penting dari kalimat yang ada di dokumen setelah dilakukan proses *case folding*. Lalu alur

proses nya ditunjukkan pada Gambar 3.5 sebagai berikut Dimana alur tersebut memvisualisasikan alur proses tahap *stopword removal*.



Gambar 3. 5 Alur proses *Stopword removal*

Pada Tabel 3.3 akan menampilkan contoh dari kalimat yang telah melalui proses *stopword removal*. Selain itu ada beberapa kata-kata yang perlu ditambahkan dalam list *stopword*, dimana kata-kata tersebut memungkinkan untuk ada pada tiap artikel berita kriminal. Kata-kata yang di-*extend* dalam list *stopword* dapat dilihat pada Tabel 3.4

Tabel 3. 3 Contoh *Filtering (Stopword removal)*

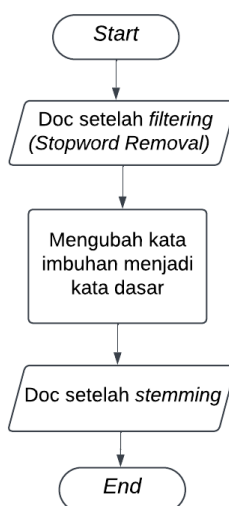
Contoh kalimat sebelum di filter	“polisi yang dilaporkan menggelapkan sembilan motor dan satu mobil warga di karimun. kepulauan riau kepri tak pernah lagi berdinis. dia sudah 10 hari tak masuk kantor.”
Contoh Kalimat sesudah proses filter dengan <i>Stopword removal</i>	“polisi laporkan menggelapkan sembilan motor satu mobil warga karimun. kepulauan riau kepri pernah lagi berdinis. sudah 10 hari masuk kantor.”

Tabel 3. 4 *List* kata tambahan *stopword*

Kata-kata yang ditambahkan pada <i>list</i> <i>stopword</i>	"polisi", "kasus", "kejahatan", "tindak", "pelaku", "lokasi", "warga", "kantor", "menurut", "ditemukan", "masyarakat", "melakukan", "terjadi", "petugas", "korban", "menyatakan", "menyebutkan", "terlibat", "melibatkan", "diduga", "mengatakan", "memiliki", "mengaku", "mengungkapkan", "penyidikan", "penangkapan", "mengakui", "menyelidiki", "mencurigai", "dilaporkan"
---	---

C. *Stemming*

Menurut Normawati & Prayogi (2021) *stemming* adalah teknik *preprocessing* teks yang melibatkan memecah kata-kata menjadi bentuk yang paling mendasar, atau kata-kata akar. *Stemming* adalah proses menghilangkan imbuhan (sufiks atau awalan) dari kata sehingga analisis teks dapat memperlakukan kata dengan akar yang sama sebagai satu kesatuan. Meskipun *varians* kata dapat dikurangi dengan *stemming*, keakuratan keluarannya tidak selalu terjamin. Proses *Stemming* dibantu dengan menggunakan *library* yaitu *sastrawi*. Diagram alur proses dari *stemming* dapat dilihat dari Gambar 3.6 sebagai berikut. Contohnya dapat dilihat pada Tabel 3.5.

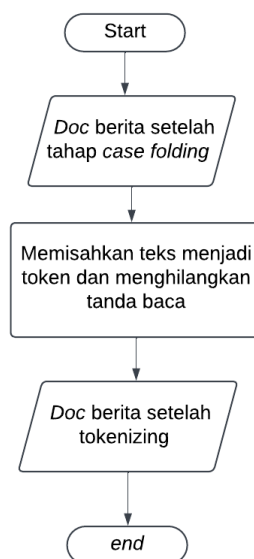
Gambar 3. 6 Alur proses *stemming*

Tabel 3. 5 Contoh hasil *Stemming*

Kata	Imbuhan	Hasil
Dilaporkan	Di (awal) dan kan (akhir)	Lapor
Menggelapkan	Meng (awal) dan kan (akhir)	Gelap
Berdinas	Ber (awal)	Dinas
Meresmikan	Me (awal) dan kan (akhir)	Resmi
Terletak	Ter (awal)	Letak

D. *Tokenizing*

Tokenizing adalah tahap dari proses *preprocessing* setelah tahapan *Stemming*. Menurut Dewi (2021) proses *tokenizing* merupakan suatu tahap untuk mengubah sebuah kalimat menjadi sebuah token atau potongan-potongan kata. Kata, kalimat, makhluk, dan karakter lain dengan makna dalam konteks teks dapat dianggap sebagai token. Tujuan mendasar tokenisasi adalah untuk membagi teks menjadi unit-unit yang lebih pasti sehingga analisis teks dapat memproses unit-unit tersebut lebih lanjut. Seperti pada proses *case folding* maupun *filtering*, tahap *tokenizing* juga menggunakan library NLTK. Untuk diagram alur proses dari *tokenizing* dapat dilihat dari Gambar 3.7 dan contoh outputnya dapat dilihat dari Tabel 3.6.

Gambar 3. 7 Alur Proses *Tokenizing*

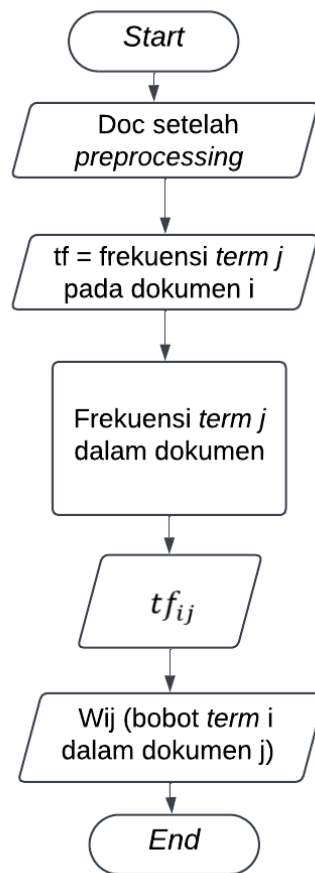
Tabel 3. 6 Contoh *Tokenizing*

Contoh kalimat sebelum <i>Tokenizing</i>	“panglima tni laksamana yudho margono kapolri jenderal listyo sigit prabowo meresmikan markas polda papua terletak kampung koya koso”
Contoh Kalimat sesudah proses <i>Tokenizing</i>	“panglima”, “tni”, “laksamana”, “yudho”, “Margono”, “kapolri”, “jenderal”, “listyo”, “sigit”, “Prabowo”, “resmi”, “markas”, “polda”, “papua”, “letak”, “kampung”, “koya”, “koso”

3.2.2 Pembobotan Kata *Term Frequency*

Pembobotan kata adalah sistem penghitungan kata berdasarkan seberapa sering setiap token muncul dalam teks atau dokumen. Pembobotan kata sering menggunakan pendekatan *Term Frequency* (TF). Metode ini menentukan frekuensi suatu kata dalam suatu dokumen (TF) (Mustofa & Mahfudh, 2019). TF adalah metode berbasis statistik yang menampilkan kemunculan suatu istilah dalam dokumen.

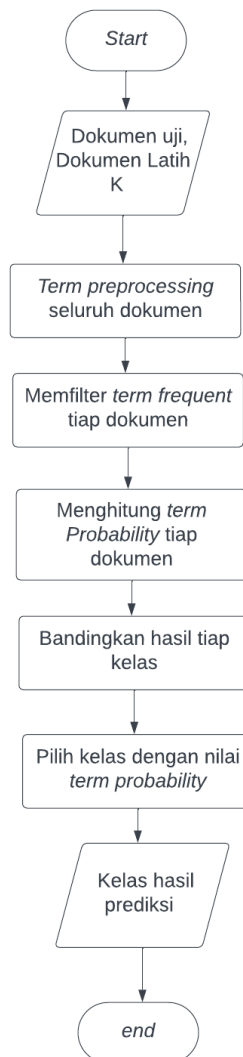
Frekuensi kemunculan sebuah kata harus dihitung untuk menentukan kata mana yang termasuk dalam sebuah kategori berdasarkan frekuensi kata yang tinggi atau rendah. tinggi atau rendah. Setiap kata dalam dokumen harus dihitung sebelum menghitung frekuensi kata. Hitungan setiap kata dalam dokumen akan disimpan dalam sebuah objek, dengan kunci adalah kata dan nilai objek adalah jumlah. Nilai bobot sebuah kata menunjukkan seberapa penting kata tersebut dalam menyajikan judul. Ketika menggunakan pembobotan TF, bobot akan lebih besar jika sebuah kata lebih sering muncul, tetapi lebih rendah jika kata tersebut lebih sering muncul di berita lainnya. Berikut adalah algoritma alur pembobotan menggunakan metode TF pada Gambar 3.8



Gambar 3. 8 Alur Proses TF

3.2.3 *Naïve Bayes Classifier*

Pada Gambar 3.9 akan menunjukkan alur proses klasifikasi menggunakan *Naïve Bayes Classifier*. Dalam proses klasifikasi *Naïve Bayes*, langkahnya melibatkan penghitungan jumlah kemunculan kata dari setiap kelas dalam *dataset* yang telah melalui beberapa tahap sebelumnya. Seperti yang dipaparkan diatas, bahwasanya metode *Naïve Bayes* merupakan metode yang mengandalkan suatu probabilitas untuk melakukan klasifikasi.



Gambar 3. 9 Alur Proses *Naïve Bayes Classifier*

Rumus umum untuk melakukan Perhitungan probabilitas dapat dilihat dari persamaan (3.1).

$$P(c|d) = \frac{P(d|c) \times P(c)}{P(d)} \quad (3.1)$$

Pada persamaan (3.2) $P(c|d)$ merupakan probabilitas prior dari dokumen d yang berada di kelas c . Adapun sebuah metode yang merupakan percabangan atau yang dikembangkan dari algoritme *Naïve Bayes* yaitu metode *Multinomial Naïve Bayes*

(MNB). MNB sangat cocok digunakan untuk melakukan proses klasifikasi pada bidang *text mining*. Dengan menggunakan MNB, jumlah kemunculan setiap kata, serta varian kata, dalam sebuah kelas dokumen akan ditentukan (Liu & Sari, 2019). Model persamaan (3.2) menggunakan MNB berbeda seperti persamaan (3.1).

$$P(c|d) = P(c) \prod_{k=1}^n P(t_k|c) \quad (3.2)$$

$P(t_k|c)$ merupakan probabilitas bersyarat *term* t_k berada di dokumen pada kelas c dan adapun $\{t_1, t_2, t_3, \dots, t_n\}$ dimana t merupakan token dalam dokumen d yang adalah bagian dari *vocabulary* dengan jumlah n . klasifikasi dokumen atau teks bertujuan untuk mengetahui adanya kemungkinan terbesar atau *maximal a posteriori* (MAP) kelas c_{map} pada persamaan (3.3)

$$c_{map} = \arg \max_{c \in C} \hat{P}(c) \prod_{k=1}^n \hat{P}(t_k|c) \quad (3.3)$$

P dinotasikan sebagai \hat{P} karena nilai sebenarnya dari $P(c|d)$ dan $P(t_k|c)$, yang akan dihitung selama proses pelatihan, tidak diketahui. Banyak probabilitas bersyarat yang dikalikan dalam Persamaan (3.3). Hal ini dapat mengakibatkan *floating point underflow*. Solusinya, menjumlahkan logaritma dari probabilitas akan meningkatkan prosedur komputasi. Kelas dengan logaritma probabilitas terbesar memiliki peluang terbaik untuk memenangkan dokumen; $\log(xy) = \log(x) + \log(y)$. Dengan persamaan (3.3) menggunakan logaritma dari probabilitas dapat dinyatakan dalam persamaan (3.4).

$$c_{map} = \arg \max_{c \in C} \left[\log \hat{P}(c) + \sum_{1 \leq k \leq n} \log \prod_{k=1}^n \hat{P}(t_k|c) \right] \quad (3.4)$$

Untuk mendapatkan nilai ($\hat{P}(c)$) menggunakan Persamaan (3.5), perlu menghitung rasio antara frekuensi dokumen ($|Doc_i|$) pada setiap kategori (c) dan

total jumlah dokumen ($|training|$) Persamaan ini memberikan probabilitas prior untuk kategori (c).

$$\hat{P}(c) = \frac{|Doc_i|}{|training|} \quad (3.5)$$

$$P(t_k|c) = \frac{n_i+1}{n+B} \quad (3.6)$$

Sementara itu, untuk mendapatkan nilai $P(t_k|c)$ dengan Persamaan (3.6), perlu menghitung rasio antara bobot TF ($n_i + 1$) pada term t_k di kategori c dan total bobot TF n seluruh term pada kategori c , ditambah dengan jumlah term dalam vocabulary B . Perhitungan ini memberikan probabilitas kondisional untuk term t_k pada kategori c . Sesuai dengan Persamaan (3.6), perhitungan nilai $P(t_k|c)$ melibatkan rasio antara bobot TF ($n_i + 1$) dari term t_k di kategori c dan total bobot TF n dari seluruh term dalam kategori c , yang kemudian ditambahkan dengan jumlah term dalam vocabulary B . Hal ini memungkinkan kita untuk memperoleh probabilitas kondisional yang esensial untuk menganalisis peran term t_k dalam konteks kategori c

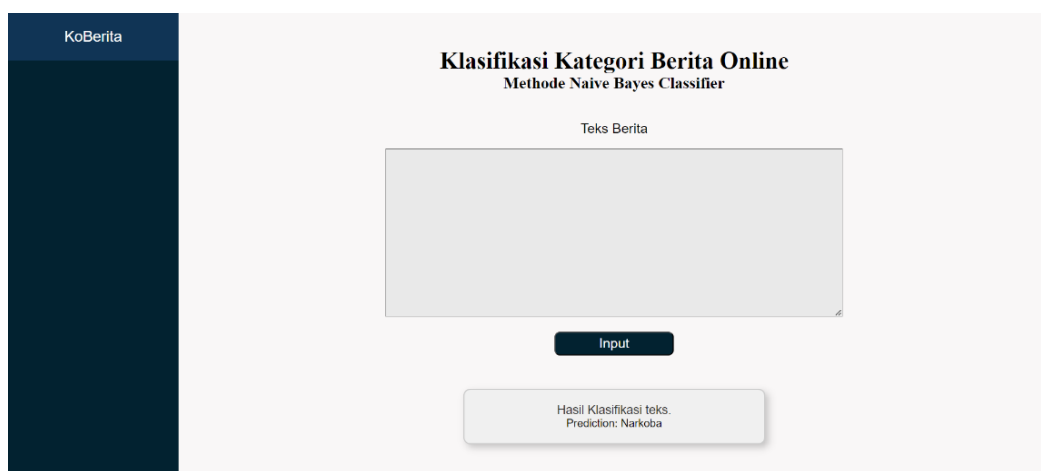
Kedua persamaan tersebut umumnya digunakan dalam konteks pemodelan klasifikasi teks, di mana ($\hat{P}(c)$) adalah probabilitas prior untuk kategori (c), dan ($P(t_k|c)$) adalah probabilitas kondisional untuk term (t_k) pada kategori

Keterangan :

$P(t_k c)$	= Probabilitas kemunculan <i>term</i> t pada suatu dokumen dengan kategori kelas c
$\hat{P}(c)$	= Probabilitas setiap dokumen pada sekumpulan dokumen
Doc_i	= frekuensi dokumen pada setiap kategori
$training$	= Jumlah dokumen <i>training</i> yang ada
n_i	= frekuensi kata ke-k pada setiap kategori
B	= jumlah kosakata atau <i>vocabulary</i> yang ada pada dokumen uji

3.3 Implementasi sistem

Sistem dirancang menggunakan *framework flask* yang mana dapat menjalankan fungsi-fungsi python. Seluruh tahap proses dari *preprocessing*, dan *modeling* menggunakan NBC dirancang hingga dapat melakukan klasifikasi berita kriminal (Kecerdasan et al., 2022). Implementasi sistem klasifikasi kategori kriminal di tampilkan pada Gambar 3.10.



Gambar 3. 10 Implementasi sistem

Pada gambar 3.10 merupakan tampilan sistem yang dibuat menjadi user interface dan juga *single-Page Application (SPA)*. *Single Page Application* merupakan suatu istilah untuk merepresentasikan tampilan yang memiliki fungsi untuk menjadi *user interface*, *input*, dan *output*. Sistem klasifikasi ini terdapat satu input utama yaitu berupa berita kriminal yang akan dikategorikan. Sistem sudah menggunakan model MNB namun sebelum input utama diprediksi akan melalui proses yang sudah dijelaskan pada sub-bab 3.2 yaitu akan melalui tahap preprocessing dulu. Tahap preprocessing dapat dilihat pada Gambar 3.11, Gambar 3.12

```

import re
import nltk
from nltk.tokenize import word_tokenize
from Sastrawi Stemmer StemmerFactory import
StemmerFactory
from nltk.corpus import stopwords

```

Gambar 3. 11 *Import Library Preprocessing*

```

def cleanTxt(text):
    text = re.sub(r'#\S+', ' ', text) # Remove Hashtags
    text = re.sub(r'\W+', '', text) # Hapus karakter
    text = re.sub(r'\w*\d+\w*', '', text) # Hapus nomor
    text = re.sub(r'\s{2,}', ' ', text) # Hapus spasi berlebih
    text = re.sub(r"^[^\w\s]", '', text) # Hapus tanda baca
    text = re.sub(r'<[<]+?>', '', text) # Hapus enter
    text = re.sub(r"\b[a-zA-Z]\b", " ", text) # Hapus karakter
    tunggal
    text = re.sub(r"#", " <hash_tag> ", text)

    return text.lower().strip() # Mengubah menjadi huruf kecil dan
    menghapus spasi di awal dan akhir

def tokenize_text(text):
    tokens = word_tokenize(text)
    return tokens

def create_stopwords():
    stopword_list = stopwords.words('indonesian')
    stopword_list.extend(["kasus", "terduga", "terjadi"])
    return stopword_list

def remove_stopwords(tokens, stopword_list):
    text_without_stopwords = [word for word in tokens if word.lower()
    not in stopword_list]
    return text_without_stopwords

def stemming_text(tokens):
    factory = StemmerFactory()
    stemmer = factory.create_stemmer()
    stemmed_tokens = [stemmer.stem(token) for token in tokens]
    return stemmed_tokens

```

Gambar 3. 12 *Proses Preprocessing*

Setelah tahap *preprocessing input* utama atau artikel berita akan langsung masuk dalam fungsi `MultinomialNaiveBayesTF()`, dimana dalam *class* tersebut artikel berita akan dihitung TF tiap katanya dan setelah bobotkan maka akan dihitung nilai probabilitas seperti yang ada pada sub-bab 3.2.3. Source code dari fungsi `MultinomialNaiveBayesTF()` dapat dilihat dari Gambar 3.13.

```
class MultinomialNaiveBayesTF:
    def __init__(self, categories):
        self.categories = categories
        self.prior_probs = {}
        self.tf_probs = {}

    def calculate_tf(self, document):
        word_counts = {word: document.count(word) for word in set(document)}
        total_words = sum(word_counts.values())
        return {word: count / total_words for word, count in
                word_counts.items()}

    def fit(self, data):
        # Hitung probabilitas prior
        total_documents = len(data)
        for category in self.categories:
            category_documents = data[data['Label'] == category]
            self.prior_probs[category] = len(category_documents) /
            total_documents

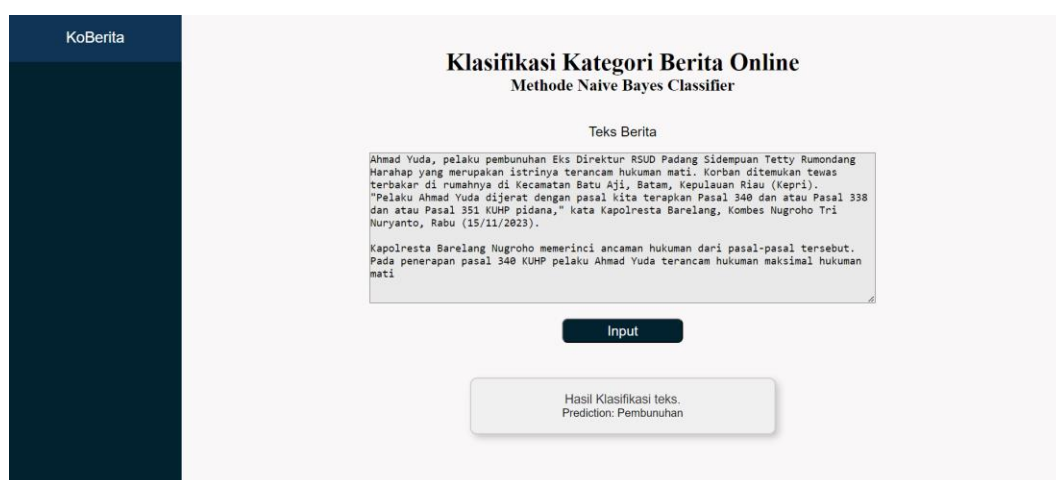
            # Hitung probabilitas term frequency
            tf_per_category = []
            for doc in category_documents['stemmer']:
                tf_per_category.append(self.calculate_tf(doc))
            self.tf_probs[category] = tf_per_category

    def predict(self, document):
        posteriors = {}
        for category in self.categories:
            prior_prob = np.log(self.prior_probs[category])
            likelihood_prob = sum([np.log(tf.get(word, 1e-10)) for word, tf in
                zip(document, self.tf_probs[category])])
            posteriors[category] = prior_prob + likelihood_prob

        return max(posteriors, key=posteriors.get)
```

Gambar 3. 13 Model *Multinomial Naïve Byes*

Setelah artikel berita dimasukkan dalam model maka akan mendapatkan nilai probabilitas dari setiap kelasnya. Kemudian untuk predictnya nilai-nilai probabilitas tersebut akan dibandingkan dan diambil nilai *max*-nya. Pada Gambar 3.14 akan menampilkan hasil prediksi dari inputan atau artikel berita yang telah melewati beberapa tahapan yang sudah dijabarkan.



Gambar 3. 14 Hasil *output*

Pada Gambar 3.14 menunjukkan hasil yang diperoleh dari klasifikasi dengan inputan artikel berita, dimana artikel berita tersebut dapat diprediksi masuk dalam kelas kategori pembunuhan. Hal ini terjadi karena nilai probabilitas yang diperoleh dari kelas pembunuhan merupakan nilai yang paling tinggi.

BAB IV

UJI COBA DAN PEMBAHASAN

4.1 Skenario Uji Coba

Penelitian ini menjalani serangkaian percobaan dengan tujuan mendapatkan beragam tingkat akurasi untuk kedua metode yang diterapkan. Keseluruhan set data dibagi menjadi dua kategori utama, yakni data pelatihan dan data pengujian, dengan variasi persentase tertentu. Setiap uji coba dilakukan pada setiap skenario pengujian sesuai dengan skenario yang telah dijelaskan secara rinci dalam Tabel 4.1.

Tabel 4. 1 Skenario uji coba

Iterasi	Jumlah <i>Data Train</i>	Jumlah <i>Data Testing</i>
1	70%	30%
2	80%	20%
3	90%	10%

Proses awal dimulai dengan mencari *ground truth*, yang berfungsi sebagai acuan untuk membandingkan hasil prediksi yang dihasilkan oleh sistem. Data yang digunakan dalam penelitian ini terdiri dari 300 artikel berita dalam bahasa Indonesia, yang terbagi ke dalam 5 kategori berita kriminal, dengan setiap kategori memiliki 60 artikel berita. *Ground truth* atau label sebenarnya yang digunakan berasal dari 5 kategori berita, yaitu perampokan, pembunuhan, pemerkosaan, korupsi, dan narkoba, sehingga artikel-artikel yang dikumpulkan dapat diklasifikasikan ke dalam salah satu dari kategori tersebut. Sumber *ground truth* dan artikel berita diperoleh dari situs web penyedia berita *online*, seperti kompas.com dan detik.com, tempat di mana semua data dikumpulkan.

Dalam fase pengujian awal, persiapannya melibatkan penggunaan 300 artikel berita. Selanjutnya, *dataset* tersebut dipecah menjadi data pelatihan dan data pengujian dengan rasio 70:30, menghasilkan 210 artikel berita untuk pelatihan dan 90 artikel berita untuk pengujian. Jumlah data untuk setiap kelas pada kedua *dataset* tersebut sudah seimbang, di mana setiap sub-kategori memiliki 42 artikel untuk pelatihan dan 18 artikel untuk pengujian. Setelah mendapatkan hasil, langkah selanjutnya adalah menghitung nilai akurasi, presisi, *recall*, dan *f-measure*.

Pada tahap pengujian berikutnya, yang sesuai dengan iterasi kedua, persiapan melibatkan penggunaan 300 artikel berita. Dalam *dataset* tersebut, dilakukan pembagian antara data pelatihan dan data pengujian dengan perbandingan 80:20, sehingga menghasilkan 240 artikel berita untuk pelatihan dan 60 artikel berita untuk pengujian. Sama halnya dengan skenario uji 80:20, di mana setiap kelas pada data pelatihan memiliki 48 artikel, dan 12 artikel untuk data pengujian. Setelah memperoleh hasil, langkah berikutnya adalah menghitung nilai akurasi, presisi, *recall*, dan *f-measure*.

Pada pengujian dalam iterasi ketiga, persiapannya melibatkan pemanfaatan 300 artikel berita. Selanjutnya, *dataset* tersebut dibagi menjadi data pelatihan dan data pengujian dengan perbandingan 90:10, sehingga menghasilkan 270 artikel berita untuk pelatihan dan 30 artikel berita untuk pengujian. Dalam iterasi ketiga ini, jumlah data pelatihan untuk setiap kelas adalah sebanyak 54 artikel, sedangkan data pengujian untuk setiap kelas adalah sebanyak 6 artikel. Setelah mendapatkan hasil, langkah selanjutnya adalah melakukan perhitungan nilai akurasi, presisi, *recall*, dan *f-measure*.

Rumus menghitung nilai dari *accuracy*, *precision*, *recall*, serta *f-measure* akan dihitung menggunakan persamaan (4.1) sampai persamaan (4.4).

$$Accuracy = \frac{TP + Tn}{TP + TN + FP + FN} \times 100\% \quad (4.1)$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (4.2)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (4.3)$$

$$f - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4.4)$$

Keterangan,

1. TP (*True Possitive*) : artikel dengan hasil klasifikasi kategori *i* yang sesuai dengan apa yang diprediksi secara akurat oleh sistem bahwa kategori *i* termasuk dalam kategori tersebut.
2. TN (*True Negative*) : artikel dengan hasil pengklasifikasian bukan kategori *i*, diprediksi benar oleh sistem bahwa kategori I tidak masuk ke kategori *i*.
3. FP (*False Positive*) : artikel dengan hasil pengklasifikasian bukan kategori *i*, namun diprediksi benar oleh sistem bahwa kategori *i* masuk ke kategori *i*.
4. FN (*False Negative*) : artikel dengan hasil pengklasifikasian kategori *i*, namun diprediksi salah oleh sistem bahwa kategori *i* tidak masuk ke kategori *i*.

4.2 Hasil Uji Coba

Evaluasi hasil pengujian dilakukan berdasarkan skenario yang telah dijelaskan sebelumnya pada sub-bab 4.1, di mana sistem akan memprediksi berita yang telah diinput dan melakukan perbandingan hasil prediksinya dengan *ground truth* atau sebutan lainnya *true label*. Pada tahap uji pertama, jumlah data yang digunakan mencapai 210 artikel sebagai data pelatihan dan 90 artikel sebagai data pengujian. Informasi terkait hasil klasifikasi untuk skenario uji 70:30 dapat ditemukan pada Tabel 4.2, sementara nilai matriks kebingungan (*confusion matrix*) terdapat pada Tabel 4.3.

Tabel 4. 2 Hasil klasifikasi pengujian 70:30

NO	True Label	Predicted Label	TP	TN	FP	FN
1	Perampokan	Perampokan	1	0	0	0
2	Perampokan	Perampokan	1	0	0	0
3	Perampokan	Pembunuhan	0	0	1	0
4	Perampokan	Pembunuhan	0	0	1	0
5	Perampokan	Perampokan	1	0	0	0
6	Perampokan	Perampokan	1	0	0	0
7	Perampokan	Perampokan	1	0	0	0
8	Perampokan	Perampokan	1	0	0	0
9	Perampokan	Perampokan	1	0	0	0
10	Perampokan	Perampokan	1	0	0	0
11	Perampokan	Perampokan	1	0	0	0
12	Perampokan	Perampokan	1	0	0	0
13	Perampokan	Perampokan	1	0	0	0
14	Perampokan	Perampokan	1	0	0	0
15	Perampokan	Perampokan	1	0	0	0
16	Perampokan	Perampokan	1	0	0	0
17	Perampokan	Perampokan	1	0	0	0
18	Perampokan	Perampokan	1	0	0	0
19	Pembunuhan	Pembunuhan	1	0	0	0
20	Pembunuhan	Pembunuhan	1	0	0	0
21	Pembunuhan	Pemerksaan	0	0	1	0
22	Pembunuhan	Narkoba	0	0	1	0
23	Pembunuhan	Pembunuhan	1	0	0	0
24	Pembunuhan	Pembunuhan	1	0	0	0
25	Pembunuhan	Pembunuhan	1	0	0	0
26	Pembunuhan	Pembunuhan	1	0	0	0
27	Pembunuhan	Pembunuhan	1	0	0	0
28	Pembunuhan	Pembunuhan	1	0	0	0
29	Pembunuhan	Pembunuhan	1	0	0	0
30	Pembunuhan	Pembunuhan	1	0	0	0
31	Pembunuhan	Pembunuhan	1	0	0	0
32	Pembunuhan	Pembunuhan	1	0	0	0
33	Pembunuhan	Pembunuhan	1	0	0	0
34	Pembunuhan	Pembunuhan	1	0	0	0
35	Pembunuhan	Pembunuhan	1	0	0	0
36	Pembunuhan	Pembunuhan	1	0	0	0
37	Pemerksaan	Pemerksaan	1	0	0	0
38	Pemerksaan	Korupsi	0	0	1	0
39	Pemerksaan	Pemerksaan	1	0	0	0
40	Pemerksaan	Pemerksaan	1	0	0	0
41	Pemerksaan	Pemerksaan	1	0	0	0
42	Pemerksaan	Pemerksaan	1	0	0	0
43	Pemerksaan	Perampokan	0	0	1	0
44	Pemerksaan	Pemerksaan	1	0	0	0
45	Pemerksaan	Pemerksaan	1	0	0	0
46	Pemerksaan	Korupsi	0	0	1	0
47	Pemerksaan	Pemerksaan	1	0	0	0
48	Pemerksaan	Korupsi	0	0	1	0
49	Pemerksaan	Pemerksaan	1	0	0	0
50	Pemerksaan	Pemerksaan	1	0	0	0
51	Pemerksaan	Pembunuhan	0	0	1	0

Lanjutan Tabel 4.2

52	Pemeriksaan	Pemeriksaan	1	0	0	0
53	Pemeriksaan	Pemeriksaan	1	0	0	0
54	Pemeriksaan	Pemeriksaan	1	0	0	0
55	Korupsi	Korupsi	1	0	0	0
56	Korupsi	Korupsi	1	0	0	0
57	Korupsi	Pembunuhan	0	0	1	0
58	Korupsi	Korupsi	1	0	0	0
59	Korupsi	Korupsi	1	0	0	0
60	Korupsi	Korupsi	1	0	0	0
61	Korupsi	Korupsi	1	0	0	0
62	Korupsi	Korupsi	1	0	0	0
63	Korupsi	Korupsi	1	0	0	0
64	Korupsi	Korupsi	1	0	0	0
65	Korupsi	Pembunuhan	0	0	1	0
66	Korupsi	Korupsi	1	0	0	0
67	Korupsi	Korupsi	1	0	0	0
68	Korupsi	Korupsi	1	0	0	0
69	Korupsi	Korupsi	1	0	0	0
70	Korupsi	Korupsi	1	0	0	0
71	Korupsi	Korupsi	1	0	0	0
72	Korupsi	Korupsi	1	0	0	0
73	Narkoba	Narkoba	1	0	0	0
74	Narkoba	Narkoba	1	0	0	0
75	Narkoba	Narkoba	1	0	0	0
76	Narkoba	Narkoba	1	0	0	0
77	Narkoba	Narkoba	1	0	0	0
78	Narkoba	Pembunuhan	0	0	1	0
79	Narkoba	Narkoba	1	0	0	0
80	Narkoba	Narkoba	1	0	0	0
81	Narkoba	Perampokan	0	0	1	0
82	Narkoba	Perampokan	0	0	1	0
83	Narkoba	Perampokan	0	0	1	0
84	Narkoba	Narkoba	1	0	0	0
85	Narkoba	Narkoba	1	0	0	0
86	Narkoba	Pembunuhan	0	0	1	0
87	Narkoba	Perampokan	0	0	1	0
88	Narkoba	Narkoba	1	0	0	0
89	Narkoba	Narkoba	1	0	0	0
90	Narkoba	Narkoba	1	0	0	0

Tabel 4. 3 *Confusion Matrix* untuk pengujian 70:30

Kategori	<i>Confusion Matrix</i>			
	TP	TN	FP	FN
Perampokan	16	67	5	2
Pembunuhan	16	65	7	2
Pemeriksaan	13	71	1	5
Korupsi	16	69	3	2
Narkoba	12	71	1	6
Total	73	343	17	17

Pengukuran akurasi, presisi, *recall*, dan *f-measure* untuk pengujian dengan rasio 70:30 dapat dijelaskan sebagai berikut:

$$Accuracy = \frac{73 + 343}{73 + 343 + 17 + 17} \times 100\% = 0.9244 \times 100\% = 92.44\%$$

$$Precision = \frac{73}{73 + 17} \times 100\% = 0.8111 \times 100\% = 81.11\%$$

$$Recall = \frac{73}{73 + 17} \times 100\% = 0.8111 \times 100\% = 81.11\%$$

$$f - measure = 2 \times \frac{81.11 \times 81.11}{81.11 + 81.11} = 2 \times 40.555 = 81.11\%$$

Perolehan akhir dari pengujian dengan pembandingan data training dan data testing 70:30 yaitu nilai *accuracy* 92.44%, *precision* 81.11%, *recall* 81.11% dan *f-measure* 81.11%.

Kemudian, pada tahap pengujian kedua, digunakan total 240 artikel berita, dengan 180 artikel sebagai data pelatihan dan 60 artikel sebagai data pengujian. Informasi hasil klasifikasi untuk pengujian dengan perbandingan 80:20 dapat ditemukan dalam Tabel 4.4, sedangkan nilai matriks kebingungan (*confusion matrix*) terdapat pada Tabel 4.5. Selanjutnya, dilakukan perhitungan untuk mendapatkan nilai akurasi, presisi, *recall*, dan *f-measure* berdasarkan matriks kebingungan pada pengujian dengan data pelatihan sebesar 80%.

Tabel 4. 4 Hasil klasifikasi pengujian 80:20

NO	True Label	Predicted Label	TP	TN	FP	FN
1	Perampokan	Perampokan	1	0	0	0
2	Perampokan	Perampokan	1	0	0	0
3	Perampokan	Perampokan	1	0	0	0
4	Perampokan	Perampokan	1	0	0	0
5	Perampokan	Perampokan	1	0	0	0
6	Perampokan	Perampokan	1	0	0	0
7	Perampokan	Perampokan	1	0	0	0
8	Perampokan	Perampokan	1	0	0	0
9	Perampokan	Perampokan	1	0	0	0
10	Perampokan	Pemeriksaan	0	0	1	0

Lanjutan Tabel 4.3

11	Perampokan	Perampokan	1	0	0	0
12	Perampokan	Perampokan	1	0	0	0
13	Pembunuhan	Pemeriksaan	0	0	1	0
14	Pembunuhan	Pembunuhan	1	0	0	0
15	Pembunuhan	Pembunuhan	1	0	0	0
16	Pembunuhan	Narkoba	0	0	1	0
17	Pembunuhan	Perampokan	0	0	1	0
18	Pembunuhan	Pemeriksaan	0	0	1	0
19	Pembunuhan	Pembunuhan	1	0	0	0
20	Pembunuhan	Pembunuhan	1	0	0	0
21	Pembunuhan	Pembunuhan	1	0	0	0
22	Pembunuhan	Pembunuhan	1	0	0	0
23	Pembunuhan	Pembunuhan	1	0	0	0
24	Pembunuhan	Pembunuhan	1	0	0	0
25	Pemeriksaan	Pemeriksaan	1	0	0	0
26	Pemeriksaan	Korupsi	0	0	1	0
27	Pemeriksaan	Pemeriksaan	1	0	0	0
28	Pemeriksaan	Pemeriksaan	1	0	0	0
29	Pemeriksaan	Pembunuhan	0	0	1	0
30	Pemeriksaan	Pemeriksaan	1	0	0	0
31	Pemeriksaan	Pembunuhan	0	0	1	0
32	Pemeriksaan	Pemeriksaan	1	0	0	0
33	Pemeriksaan	Pembunuhan	0	0	1	0
34	Pemeriksaan	Pemeriksaan	1	0	0	0
35	Pemeriksaan	Pembunuhan	0	0	1	0
36	Pemeriksaan	Pemeriksaan	1	0	0	0
37	Korupsi	Korupsi	1	0	0	0
38	Korupsi	Korupsi	1	0	0	0
39	Korupsi	Korupsi	1	0	0	0
40	Korupsi	Korupsi	1	0	0	0
41	Korupsi	Korupsi	1	0	0	0
42	Korupsi	Korupsi	1	0	0	0
43	Korupsi	Korupsi	1	0	0	0
44	Korupsi	Korupsi	1	0	0	0
45	Korupsi	Korupsi	1	0	0	0
46	Korupsi	Korupsi	1	0	0	0
47	Korupsi	Korupsi	1	0	0	0
48	Korupsi	Korupsi	1	0	0	0
49	Narkoba	Narkoba	1	0	0	0
50	Narkoba	Narkoba	1	0	0	0
51	Narkoba	Narkoba	1	0	0	0
52	Narkoba	Narkoba	1	0	0	0
53	Narkoba	Perampokan	0	0	1	0
54	Narkoba	Narkoba	1	0	0	0
55	Narkoba	Narkoba	1	0	0	0
56	Narkoba	Pemeriksaan	0	0	1	0
57	Narkoba	Narkoba	1	0	0	0
58	Narkoba	Narkoba	1	0	0	0
59	Narkoba	Pembunuhan	0	0	1	0
60	Narkoba	Narkoba	1	0	0	0

Tabel 4. 5 *Confusion Matrix* untuk pengujian 80:20

Kategori	<i>Confusion Matrix</i>			
	TP	TN	FP	FN
Perampokan	11	46	2	1
Pembunuhan	8	43	5	4
Pemeriksaan	7	44	4	5
Korupsi	12	47	1	0
Narkoba	9	47	1	3
Total	47	227	13	13

$$Accuracy = \frac{47 + 227}{47 + 227 + 13 + 13} \times 100\% = 0.9133 \times 100\% = 91.33\%$$

$$Precision = \frac{47}{47 + 13} \times 100\% = 0.7833 \times 100\% = 78.33\%$$

$$Recall = \frac{47}{47 + 13} \times 100\% = 0.7833 \times 100\% = 78.33\%$$

$$f - measure = 2 \times \frac{78.33 \times 78.33}{78.33 + 78.33} = 2 \times 39.165 = 78.33\%$$

Hasil akhir dari pengujian dengan rasio 80:20 adalah, *accuracy* 91.33%, *precision* 78.33%, *recall* 78.33%, dan *f-measure* 78.33%,

Dalam tahap pengujian selanjutnya, yang merupakan skenario ketiga, *dataset* dibagi menjadi data pelatihan dan data pengujian dengan perbandingan 90:10. Sejumlah 270 artikel berita digunakan sebagai data pelatihan, sementara 30 artikel berita digunakan sebagai data pengujian. Hasil klasifikasi untuk pengujian dengan rasio 90:10 tercatat dalam Tabel 4.6, dan nilai matriks kebingungan (*confusion matrix*) dapat ditemukan pada Tabel 4.7. Setelah itu, dilakukan perhitungan akurasi, presisi, recall, dan *f-measure* berdasarkan *confusion matrix*.

Tabel 4. 6 Hasil klasifikasi pengujian 90:10

NO	True Label	Predicted Label	TP	TN	FP	FN
1	Perampokan	Pembunuhan	0	0	1	0
2	Perampokan	Perampokan	1	0	0	0
3	Perampokan	Perampokan	1	0	0	0
4	Perampokan	Perampokan	1	0	0	0

Lanjutan Tabel 4.6

5	Perampokan	Perampokan	1	0	0	0
6	Perampokan	Perampokan	1	0	0	0
7	Pembunuhan	Pembunuhan	1	0	0	0
8	Pembunuhan	Pembunuhan	1	0	0	0
9	Pembunuhan	Perampokan	0	0	1	0
10	Pembunuhan	Pembunuhan	1	0	0	0
11	Pembunuhan	Pembunuhan	1	0	0	0
12	Pembunuhan	Pembunuhan	1	0	0	0
13	Pemeriksaan	Pemeriksaan	1	0	0	0
14	Pemeriksaan	Pemeriksaan	1	0	0	0
15	Pemeriksaan	Pemeriksaan	1	0	0	0
16	Pemeriksaan	Pemeriksaan	1	0	0	0
17	Pemeriksaan	Pemeriksaan	1	0	0	0
18	Pemeriksaan	Pembunuhan	0	0	1	0
19	Korupsi	Korupsi	1	0	0	0
20	Korupsi	Korupsi	1	0	0	0
21	Korupsi	Perampokan	0	0	1	0
22	Korupsi	Korupsi	1	0	0	0
23	Korupsi	Korupsi	1	0	0	0
24	Korupsi	Korupsi	1	0	0	0
25	Narkoba	Perampokan	0	0	1	0
26	Narkoba	Narkoba	1	0	0	0
27	Narkoba	Narkoba	1	0	0	0
28	Narkoba	Narkoba	1	0	0	0
29	Narkoba	Narkoba	1	0	0	0
30	Narkoba	Pembunuhan	0	0	1	0

Tabel 4.7 *Confusion Matrix* untuk pengujian 90:10

Kategori	<i>Confusion Matrix</i>			
	TP	TN	FP	FN
Perampokan	5	21	3	1
Pembunuhan	5	21	3	1
Pemeriksaan	5	24	0	1
Korupsi	5	24	0	1
Narkoba	4	24	0	2
Total	24	114	6	6

$$Accuracy = \frac{24 + 114}{24 + 114 + 6 + 6} \times 100\% = 0.92 \times 100\% = 92\%$$

$$Precision = \frac{24}{24 + 6} \times 100\% = 0.8 \times 100\% = 80\%$$

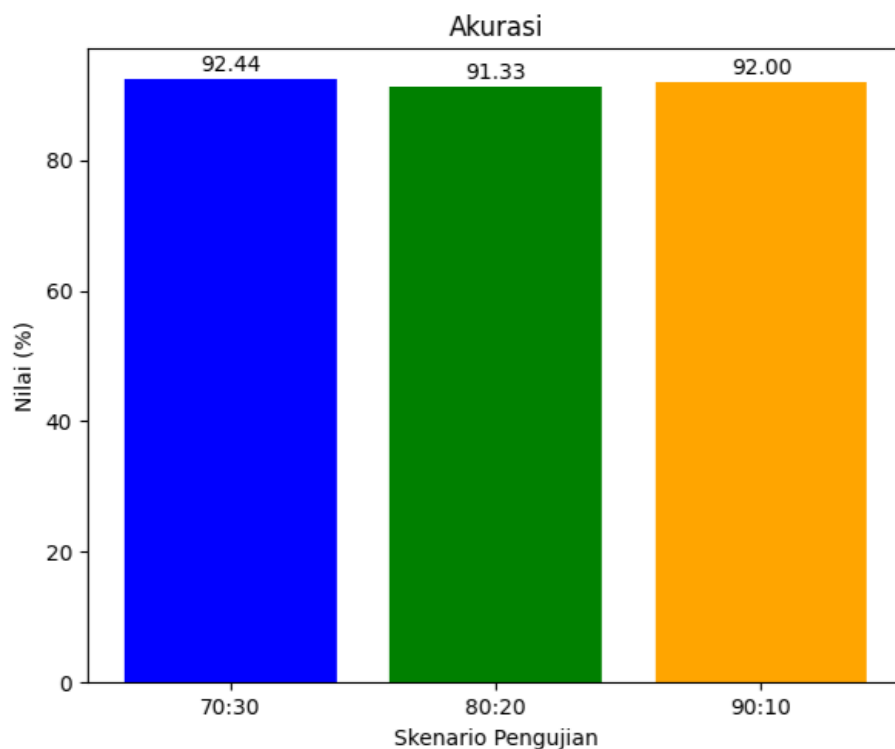
$$Recall = \frac{24}{24 + 6} \times 100\% = 0.8 \times 100\% = 80\%$$

$$f - measure = 2 \times \frac{80 \times 80}{80 + 80} = 2 \times 40 = 80\%$$

Hasil akhir dari pengujian dengan perbandingan 90:10 *accuracy* 92%, *precision* 80%, *recall* 80%, dan *f-measure* 80%.

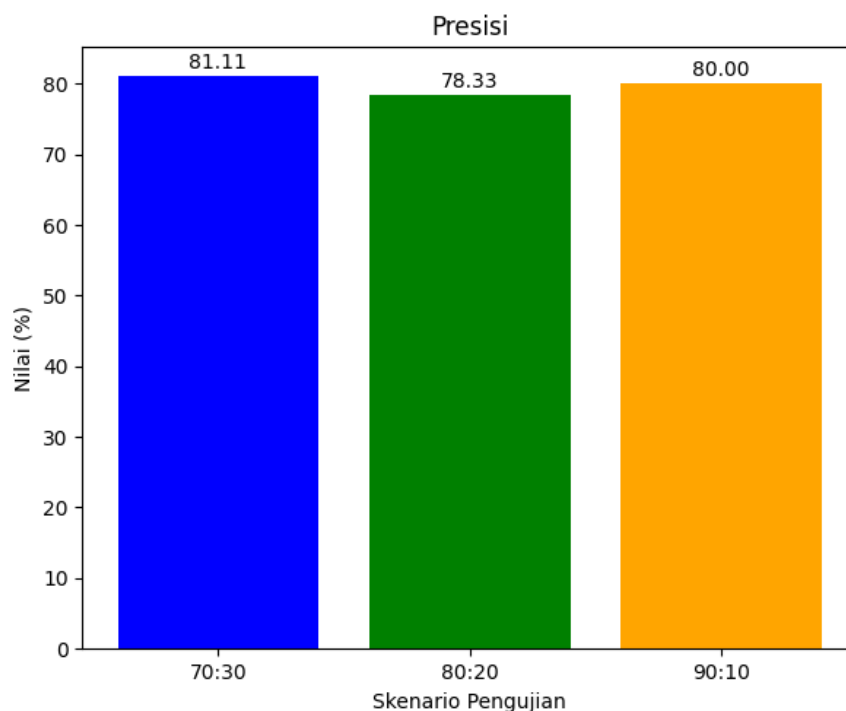
4.3 Pembahasan

Hasil akhir pengujian dengan rasio 70:30, dihitung dengan menggunakan Persamaan 4.1, menunjukkan tingkat akurasi sebesar 92.44%, presisi 81.11%, *recall* 81.11%, dan *f-measure* 81.11%. Sementara itu, pengujian dengan rasio 80:20 mencapai akurasi 91.33%, presisi 78.33%, *recall* 78.33%, dan *f-measure* 78.33%. Pengujian 90:10 memberikan akurasi 92%, presisi 80%, *recall* 80%, dan *f-measure* 80%. Selain itu, hasil yang konsisten pada berbagai rasio pembagian data menunjukkan kestabilan model dalam mengatasi variasi *dataset*, mendukung keandalan hasil pengujian secara umum. Dapat dilihat perbandingan untuk akurasi ketiga pengujian pada Gambar 4.1.



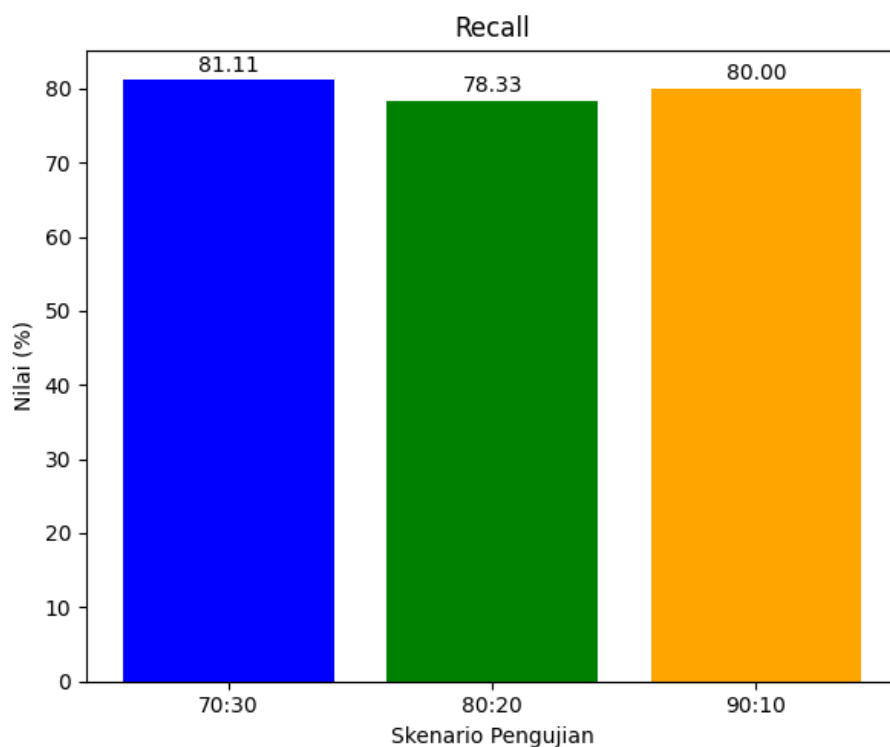
Gambar 4. 1 Grafik perbandingan hasil akurasi ketiga pengujian

Pada gambar tersebut nilai akurasi tertinggi dihasilkan pada skenario pengujian dengan rasio 70:30. Hasil tersebut menyatakan bahwa model MNB cenderung memberikan prediksi yang benar. Akurasi yang tinggi menandakan kinerja model yang baik dalam mengklasifikasikan data baru. Terjadi penurunan akurasi pada pengujian 80:20, namun tingkat akurasi yang dihasilkan model masih tinggi yaitu sebesar 91.33%. Dan pengujian 90:10 menghasilkan akurasi yang cukup dekat dengan skenario 70:30, hal ini menunjukkan bahwa model tidak mengalami *overfitting* bahkan dengan data pelatihan yang digunakan lebih banyak. Dalam konteks pengujian model, perlu juga diperhatikan faktor-faktor seperti kecepatan pelatihan dan kompleksitas model agar hasil evaluasi dapat memberikan gambaran yang holistik terkait performa model. Visualisasi hasil presisi untuk ketiga pengujian dapat dilihat pada Gambar 4.2.



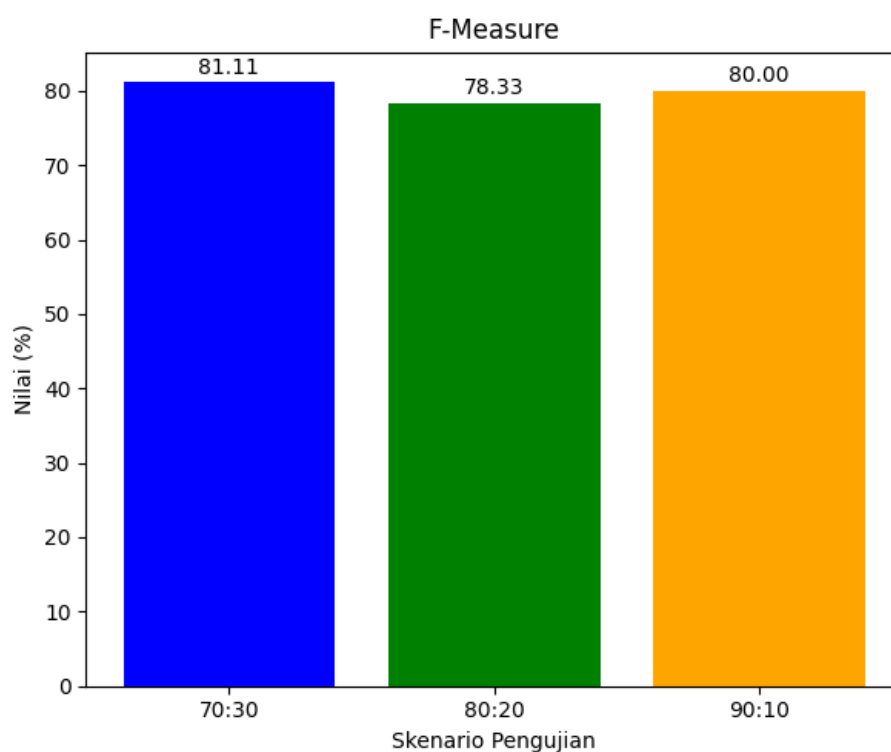
Gambar 4. 2 Grafik perbandingan hasil presisi ketiga pengujian

Hasil presisi ketiga pengujian tersebut menunjukkan bahwa model cenderung memberikan prediksi positif yang akurat, dan tidak memberikan banyak FP, terlepas dari variasi dalam pembagian data pelatihan dan pengujian. Walaupun ada terjadi penurunan presisi yang terjadi dari 70:30 ke pengujian selanjutnya. Namun, hasil presisi dari skenario 80:20 dan 90:20 masih relatif tinggi menandakan model memiliki kemampuan yang baik. Dan seperti yang disebutkan sebelumnya, keputusan akhir tentang pemisahan data optimal perlu mempertimbangkan faktor-faktor tambahan dan dapat memerlukan eksperimen lebih lanjut. Dengan demikian, perlu dilakukan evaluasi lebih lanjut untuk memahami faktor-faktor yang mempengaruhi penurunan presisi dan untuk menentukan strategi optimal guna meningkatkan performa model pada skenario pengujian yang lebih berat.



Gambar 4. 3 Grafik perbandingan hasil *recall* ketiga pengujian

Berdasarkan Gambar 4.3 menunjukkan hasil model untuk mengukur sejauh mana dapat mengidentifikasi semua *instance positive* yang sebenarnya pada ketiga pengujian. Dalam hal mengidentifikasi kasus positif yang sebenarnya, model melakukan pekerjaan yang baik dan tidak terlalu sering melewatkan kasus positif. Meskipun ada variasi dalam nilai *recall* antara skenario, hasilnya masih memadai dan menunjukkan konsistensi model dalam menangkap kasus positif yang sebenarnya. Namun, perlu diperhatikan bahwa terdapat perbedaan kinerja *recall* antara skenario pengujian, yang mungkin memerlukan penyesuaian model untuk meningkatkan konsistensinya dalam mengidentifikasi semua *instance* positif yang sebenarnya pada setiap uji coba.. Dan untuk visualisasi grafik *f-measure* dapat dilihat pada Gambar 4.4.



Gambar 4. 4 Grafik perbandingan hasil *F-Measure* ketiga pengujian

F-measure memberikan gambaran keseluruhan tentang kinerja model dalam menyeimbangkan antara jumlah positif yang benar dan jumlah positif yang sebenarnya. Pada grafik menggambarkan model memiliki keseimbangan yang baik antara akurasi dalam mengidentifikasi kasus positif dan kemampuan menangkap sebagian besar kasus positif yang sebenarnya. Ini menunjukkan juga bahwa model memiliki kinerja yang solid dalam tiga skenario pengujian, dan pilihan pemisahan data untuk pelatihan dan pengujian tidak secara signifikan memengaruhi keseimbangan antara presisi dan *recall*.

Pada saat pengujian dapat disimpulkan secara garis besar bahwa penggunaan 3 skenario pengujian dengan rasio yang berbeda beda tetap menghasilkan nilai akurasi yang konsisten. Jika dilihat dari pada hasil uji coba, terdapat perbedaan dalam performa MNB kategorisasi ketika membandingkan tiga skenario pengujian yaitu 70:30, 80:20, dan 90:10. Model MNB ini memiliki kinerja yang relatif lebih baik pada skenario 70:30, di mana 70% data digunakan untuk pelatihan, dengan akurasi, presisi, *recall*, dan *f-measure* yang lebih tinggi. Peningkatan *volume* data pelatihan dapat menjelaskan hasil ini dengan memberikan lebih banyak informasi kepada model untuk mengidentifikasi pola-pola dalam data. Namun, terjadi penurunan performa model pada situasi 80:20, di mana hanya 80% data yang digunakan untuk pelatihan. Penurunan ini dapat dikaitkan dengan model yang kurang mewakili variabel-variabel dalam *dataset* secara lengkap, sehingga kurang mampu menangkap informasi penting selama pengujian.

Sementara itu, dalam situasi 90:10, di mana 90% dari data digunakan untuk pelatihan, model ini bekerja dengan sangat baik dan akurat. Namun, perlu

ditekankan bahwa ukuran *dataset* yang berkurang dapat berdampak pada hasil ini, dan bahaya *overtraining* harus dipertimbangkan. Secara keseluruhan, perbandingan dari ketiga kasus tersebut menekankan pada sulitnya menentukan pembagian data terbaik. *Volume* data pelatihan, distribusi kelas, dan ketidakpastian hasil eksperimen semuanya berdampak pada kinerja dari menggunakan metode *Naïve Bayes Classifier*. Penyaringan lebih lanjut terhadap karakteristik *dataset* dan variabilitas hasil dapat memberikan lebih banyak wawasan tentang variabel yang mempengaruhi variasi kinerja di ketiga skenario.

Pengujian sistem klasifikasi pada tiga skenario pembagian data yang berbeda memberikan gambaran lengkap mengenai dampaknya terhadap kinerja model MNB. Hasilnya menunjukkan bahwa fraksi data pelatihan yang dipilih memiliki dampak yang signifikan terhadap kapasitas model untuk mengenali pola dalam kumpulan data. Manfaat dari pengujian ini adalah memberikan pemahaman yang lebih baik tentang bagaimana *volume* data pelatihan mempengaruhi pengembangan model yang efektif. Hasil yang lebih baik dalam skenario dengan persentase pelatihan yang lebih besar menyoroti relevansi memiliki *dataset* yang mencakup berbagai kondisi dan varian untuk meningkatkan generalisasi model. Sistem klasifikasi diharapkan mampu menyajikan solusi yang lebih efisien dan dapat diandalkan dalam proses klasifikasi berita kriminal.

Pada topik pembahasan lebih menjelaskan hasil akurasi, presisi, *recall*, dan *f-measure* dari klasifikasi berita kriminal dimana dalam Islam ada suatu hadist yang berasal dari koleksi hadits Sahih al-Bukhari, Hadits ini termasuk dalam Kitab al-Adab (Buku tentang Etika) di dalam Sahih al-Bukhari.

عَنْ أَبِي هُرَيْرَةَ، قَالَ قَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ " مَنْ كَانَ يُؤْمِنُ بِاللَّهِ وَالْيَوْمِ الْآخِرِ فَلْيُكَلِّمْ خَيْرًا
أَوْ لِيُصْمُتْ

Dari Abu Hurairah, ia berkata: Rasulullah SAW bersabda, "Barangsiapa yang beriman kepada Allah dan hari kiamat, hendaklah ia berkata yang baik atau diam." (Sahih al-Bukhari, Kitab al-Adab, Bab Qawl al-Tayyib Awr al-Samt)

Pada hadist menekankan beberapa hal, Seseorang yang berada dalam posisi untuk memberikan komentar, mengungkapkan pendapat, berdebat atau diskusi, ataupun memberikan saran dan nasehat. Dalam konteks ini, akurasi dapat disamakan dengan sejauh mana seseorang menyampaikan pendapat yang baik dan benar sesuai dengan keyakinan Islam, atau, jika tidak mampu, lebih baik diam untuk menghindari perkataan yang tidak baik. Pada konteks presisi dapat diibaratkan sebagai kemampuan seseorang untuk memberikan saran atau nasihat yang sesuai dan tepat, tanpa menyertakan informasi yang tidak relevan atau salah. Dalam konteks lain yaitu *recall* diibaratkan sebagai sejauh mana orang tersebut dapat mengingat dan menerapkan pelajaran dari hadist, misalnya, untuk berbicara dengan baik atau diam, dalam kasus-kasus seperti itu. Sedangkan untuk konteks *f-measure* diibaratkan sebagai sejauh mana individu mencapai keseimbangan antara memberikan kontribusi yang berguna (ketepatan) dan memahami kapan harus diam untuk mencegah perselisihan atau kata-kata yang tidak baik (daya ingat).

Dalam konteks klasifikasi berita online, implementasi *mu'amalah ma'a Allah* dalam penelitian ini mencakup pemberian manusia kekuatan untuk mengembangkan teknologi yang bertujuan mempermudah pelaksanaan tugas-tugas mereka. Penelitian ini diyakini akan menjadi kontribusi berarti bagi masyarakat dalam memahami realitas saat ini. Dengan membantu redaksi dalam pekerjaan

mereka dan memudahkan masyarakat dalam mendapatkan informasi berita, penelitian ini diharapkan dapat menginspirasi masyarakat untuk selalu mengingat Allah Subhanahu wa ta'ala dan bersyukur atas segala karunia-Nya. Pemahaman dan penerapan *mu'amalah ma'a Allah* dalam konteks teknologi memberikan dimensi spiritual yang lebih mendalam, memandang bahwa penggunaan teknologi dapat menjadi sarana untuk mendekatkan diri kepada-Nya dan menguatkan hubungan dengan sesama.

Sebagai analogi, dapat merujuk pada Qur'an Surah An-Nisa ayat 135 Allah Subhanahu wa ta'ala berfirman:

يَا أَيُّهَا الَّذِينَ آمَنُوا كُونُوا قَوَّامِينَ بِالْقِسْطِ شُهَدَاءَ لِلَّهِ وَلَوْ عَلَىٰ أَنفُسِكُمْ أَوِ الْوَالِدِينَ وَالْأَقْرَبِينَ ۚ إِن يَكُنْ
عَنِيًّا أَوْ فَقِيرًا فَاللَّهُ أَوْلَىٰ بِمِمَّا فَلَآ تَتَّبِعُوا أَهْوَىٰ ۚ إِن تَعَدِلُوا ۚ وَإِن تَلْوُوا ۚ أَوْ تُعْرِضُوا فَإِنَّ اللَّهَ كَانَ بِمَا تَعْمَلُونَ خَبِيرًا

“Wahai orang-orang yang beriman! Jadilah kamu penegak keadilan, menjadi saksi karena Allah, walaupun terhadap dirimu sendiri atau terhadap ibu bapak dan kaum kerabatmu. Jika dia (yang terdakwa) kaya ataupun miskin, maka Allah lebih tahu kemaslahatan (kebaikannya). Maka janganlah kamu mengikuti hawa nafsu karena ingin menyimpang dari kebenaran. Dan jika kamu memutarbalikkan (kata-kata) atau enggan menjadi saksi, maka ketahuilah Allah Mahateliti terhadap segala apa yang kamu kerjakan.” (QS An-Nisa:135)

Al-Qur'an Surat An-Nisa (4:135) memberikan informasi penting dalam konteks muammalah (urusan sosial dan transaksi manusia) yang dapat dikaitkan dengan pengembangan sistem klasifikasi kategori kriminal. Ayat ini mendorong orang beriman untuk menjadi orang yang teguh dalam menegakkan keadilan, bahkan jika itu berarti memberikan kesaksian terhadap diri sendiri, keluarga, dan kerabat. Pesan ini penting ketika mengembangkan sistem klasifikasi yang adil dan sukses. Anjuran keadilan dalam surah ini dapat dipandang sebagai landasan moral

untuk pengembangan sistem klasifikasi. Dalam klasifikasi kriminal, keadilan membutuhkan perlakuan yang adil terhadap semua individu, tanpa memandang status sosial atau pendapatan. Berdasarkan nilai-nilai Islam, sistem klasifikasi ini harus melambangkan keadilan.

Kemudian, implementasi *mu'amalah ma'a An-nas* dalam penelitian ini terutama ditujukan untuk mempermudah tugas editor dalam mengklasifikasikan berita kriminal. Dengan demikian, editor tidak perlu lagi melakukan proses klasifikasi secara manual. Hal ini berdampak pada kemudahan akses pembaca terhadap berita yang ingin mereka baca, dengan memungkinkan mereka menemukan berita sesuai preferensi secara cepat. Pembaca dapat dengan mudah menelusuri dan membaca berita sesuai dengan kategori yang diinginkan tanpa khawatir apakah berita tersebut terdapat pada tempat yang tepat. Hasilnya, penelitian ini diyakini akan memudahkan editor untuk mengklasifikasikan berita. Editor membantu publik menemukan berita dengan mengkategorikan dan mengklasifikasikannya. Dalam Al-Qur'an surah Al-Hujurat Allah Subhanahu wa ta'ala berfirman:

يَا أَيُّهَا الَّذِينَ آمَنُوا إِنْ جَاءَكُمْ فَاسِقٌ بِنَبَأٍ فَتَبَيَّنُوا أَنْ تُصِيبُوا قَوْمًا بِجَهَالَةٍ فَتُصْحِحُوا عَلَىٰ مَا فَعَلْتُمْ نُدْمِينَ

“Hai orang-orang yang beriman, jika datang kepadamu orang fasik membawa suatu berita, maka periksalah dengan teliti, agar kamu tidak menimpakan suatu musibah kepada suatu kaum tanpa mengetahui keadaannya yang menyebabkan kamu menyesal atas perbuatanmu itu“. (QS. Al-Hujurat 49:6)

Dalam tafsir Al-Muyyasar (TafsirWeb, tanpa tanggal), dijelaskan bahwa bagi mereka yang teguh dalam keimanannya kepada Allah dan Rasul-Nya serta mentaati peraturan-peraturan-Nya, apabila ada individu yang tidak bermoral

membawa suatu berita, disarankan untuk melakukan pengecekan terlebih dahulu sebelum mempercayai dan menyebarkannya. Hal ini dimaksudkan agar kita dapat memastikan kebenaran informasi tersebut. Dengan melakukan hal ini, dihindari kemungkinan berbuat zalim terhadap suatu kelompok yang tidak bersalah, yang kemudian dapat menimbulkan penyesalan di kemudian hari.

Dalam ayat tersebut, Allah menjelaskan, "Hai orang-orang yang beriman, jika datang kepada kalian orang fasik membawa suatu berita, maka periksalah dengan teliti." Pesannya adalah bahwa ketika kita menerima suatu informasi, kita seharusnya melakukan penyelidikan lebih lanjut untuk memastikan kebenaran dari informasi tersebut. Diharapkan bahwa sistem klasifikasi berita kriminal ini akan memberikan manfaat bagi para editor dalam menjalankan tugas mereka, yaitu mengelompokkan berita-berita ke dalam kategori yang sesuai. Hal ini diharapkan dapat memudahkan masyarakat untuk mengakses berita berdasarkan kategori yang diminati dan sekaligus meningkatkan kewaspadaan terhadap tindak kriminal. Dengan demikian, upaya ini diharapkan dapat meningkatkan keakuratan dan kebermanfaatan informasi yang disajikan kepada masyarakat.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

1. Hasil pengujian dari ketiga skenario adalah, Skenario 70:30 memiliki performa terbaik, dengan akurasi 92,44%, presisi 81,11%, recall 81,11%, dan f-measure 81,11%. Terjadi penurunan performa pada skenario 80:20, dengan akurasi 91,33%, presisi 78,33%, recall 78,33%, dan f-measure 78,33%. dan Skenario 90:10 menghasilkan akurasi 92%, presisi 80%, recall 80%, dan f-measure 80%.
2. Ada beberapa faktor yang dapat mempengaruhi klasifikasi dimana meningkatkan jumlah data pelatihan, seperti pada skenario pengujian 70:30, memberikan informasi tambahan kepada model, meningkatkan kapasitasnya untuk menemukan pola dalam data. Sedangkan penurunan *volume* data, seperti pada skenario 80:20, dapat mengakibatkan kurang terwakili variabel dalam *dataset*, sehingga mempengaruhi kinerja model. Pembagian data yang ideal sulit untuk diidentifikasi, karena pilihannya dipengaruhi oleh *volume* data pelatihan, dan ketidakpastian hasil eksperimen.

5.2 Saran

1. Sangat menguntungkan apabila terdapat *volume* data yang lebih besar, karena hal tersebut dapat berdampak positif pada performa sistem, sehingga prediksi yang dihasilkan memiliki tingkat akurasi yang lebih tinggi.
2. Dengan melakukan penambahan fitur atau teknik *preprocessing* dapat dievaluasi untuk meningkatkan kinerja model. Penggunaan kata-kata kunci, *n-*

gram, dan penyelidikan fitur ekstraksi informasi dapat membantu meningkatkan representasi teks.

3. Dapat melakukan klasifikasi berita kriminal yang memiliki lebih dari satu sub-kategori atau artikel *multiple crime*, karena tidak semua berita kriminal hanya memiliki satu sub-kategori.

DAFTAR PUSTAKA

- Al Fathan, R. A., & Aminudin, A. (2021). Strategi Tim Visual Indozone.Id Dalam Perancangan Infografik Di Media Sosial Instagram. *Jurnal Kajian Media*, 5(1), 60–75. <https://doi.org/10.25139/jkm.v5i1.3625>
- Alita, D., & Isnain, A. R. (2020). Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier. *Jurnal Komputasi*, 8(2), 50–58. <https://doi.org/10.23960/komputasi.v8i2.2615>
- Anshori, M. A. I. (2020). Perbandingan Metode Naïve Bayes Dengan K-Nearest Neighbor (KNN) Untuk Klasifikasi Kategori Abstrak Skripsi. *Tidak Diterbitkan*. <http://etheses.uin-malang.ac.id/id/eprint/18148>
- Arsi, P., Prayoga, I., & Asyari, M. H. (2023). Klasifikasi Sentimen Publik Terhadap Jenis Vaksin Covid-19 yang Tersertifikasi WHO Berbasis NLP dan KNN. *Jurnal Media Informatika Budidarma*, 7(1), 260–266. <https://doi.org/10.30865/mib.v7i1.5418>
- Cucus, A., Endra, R. Y., & Neralita, T. (2019). Chatter Bot Untuk Konsultasi Akademik Di Perguruan Tinggi. *Explore: Jurnal Sistem Informasi Dan Telematika*, 10(1). <https://doi.org/10.36448/jsit.v10i1.1214>
- Dewi, F. K. S. (2021). Klasifikasi Berita Menggunakan Metode Multinomial Naïve Bayes. *SCAN - Jurnal Teknologi Informasi Dan Komunikasi*, 16(3), 1–8. <https://doi.org/10.33005/scan.v16i3.2870>
- Habib, S. M., Haerani, E., Gusti, S. K., & Ramadhani, S. (2022). Klasifikasi Berita Menggunakan Metode Naïve Bayes Classifier. *Jurnal Nasional Komputasi Dan Teknologi Informasi*, 5(2), 248–258. <https://doi.org/10.32672/jnkti.v5i2.4191>
- Haris, A. R., Putra, G. A. W., Suartana, I. P. E., & ... (2022). Klasifikasi Warna Kulit Pada Sebuah Citra Digital Menggunakan Metode Naive Bayes. ... *Ilmu Komputer Dan ...*, 805–811. <https://conference.upnvj.ac.id/index.php/senamika/article/view/2262%0Ahttps://conference.upnvj.ac.id/index.php/senamika/article/download/2262/1776>
- Indrayuni, E. (2019). Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes. *Jurnal Khatulistiwa Informatika*, 7(1), 29–36. <https://doi.org/10.31294/jki.v7i1.1>
- Juliana, N. E., Khansa, F. D., Azis, A. M. H., Gunawan, R. I., & Cahya, N. D. (2021). Klasifikasi Kategori Berita menggunakan Algoritma Support Vector Machine. *Gunung Djati Convergence Series*, 3, 62–67.

- Kecerdasan, J., Informasi, T., Firdaus, D. H., Imran, B., Bakti, L. D., & Suryadi, E. (2022). Klasifikasi Penyakit Katarak Pada Mata Menggunakan Metode Convolutional Neural Network (Cnn) Berbasis Web Web-Based Classification of Cataract in the Eyes Using Convolutional Neural Network (Cnn) Method. *Jurnal Kecerdasan Buatan Dan Teknologi Informasi (JKBTI)*, 1(3), 18–26.
- KOMINFO. (2022). *Kominfo Lanjutkan Lima Program Prioritas di 2023*. Kementerian Komunikasi Dan Informatika. <https://www.kominfo.go.id/content/detail/44678/kominfo-lanjutkan-lima-program-prioritas-di-2023/0/artikel>
- Liu, I., & Sari, Y. A. (2019). Klasifikasi Hate Speech Berbahasa Indonesia di Twitter Menggunakan Naive Bayes dan Seleksi Fitur Information Gain dengan Normalisasi Kata. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(5), 4914–4922.
- Mandar, G., Muhamamd, A. H., & Sudin, S. (2020). Klasifikasi Berita Indonesia Menggunakan Naive Bayes dengan Porter Stemmer. *Jurnal Teknik Informatika (J-Tifa)*, 3(2), 17–22. <https://doi.org/10.52046/j-tifa.v3i2.1121>
- Matthew Turland. (2010). *Php-Architect's Guide to Web Scraping*. Marco Tabini & Associates, Incorporated.
- Megadiningrum, R., Hermawati, A., & Noviekayanti, I. (2022). Sistem Pakar Konseling Untuk Mengurangi Trauma Anak Yatim Piatu Pada Masa Pandemi Covid 19 Berbasis Natural Language Processing. *Prosiding Senakama*, 1(September), 675–685.
- Mustofa, H., & Mahfudh, A. A. (2019). Klasifikasi Berita Hoax Dengan Menggunakan Metode Naive Bayes. *Walisongo Journal of Information Technology*, 1(1), 1. <https://doi.org/10.21580/wjit.2019.1.1.3915>
- Mutawalli, L., Zaen, M. T. A., & Bagye, W. (2019). KLASIFIKASI TEKS SOSIAL MEDIA TWITTER MENGGUNAKAN SUPPORT VECTOR MACHINE (Studi Kasus Penusukan Wiranto). *Jurnal Informatika Dan Rekayasa Elektronik*, 2(2), 43. <https://doi.org/10.36595/jire.v2i2.117>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naive Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711.
- Nugroho, K. S., Istiadi, I., & Marisa, F. (2020). Naive Bayes classifier optimization for text classification on e-government using particle swarm optimization. *Jurnal Teknologi Dan Sistem Komputer*, 8(1), 21–26. <https://doi.org/10.14710/jtsiskom.8.1.2020.21-26>

- Prasetyo, V. R., Benarkah, N., & Chrisintha, V. J. (2021). Implementasi Natural Language Processing Dalam Pembuatan Chatbot Pada Program Information Technology Universitas Surabaya. *Teknika*, 10(2), 114–121. <https://doi.org/10.34148/teknika.v10i2.370>
- Putra, Z. P., & Nugroho, A. (2021). Pebandingan Performa Naïve Bayes dan KNN pada Klasifikasi Teks Sentimen Jasa Ekspedisi. *JOINTECS (Journal of Information Technology and Computer Science)*, 6(3), 145. <https://doi.org/10.31328/jointecs.v6i3.2635>
- Rhomadhona, H., & Permadi, J. (2019). Klasifikasi Berita Kriminal Menggunakan Naïve Bayes Classifier (NBC) dengan Pengujian K-Fold Cross Validation. *Jurnal Sains Dan Informatika*, 5(2), 108–117. <https://doi.org/10.34128/jsi.v5i2.177>
- Ruus, E. G., Latumakulita, L. A., & ... (2022). Analisis Sentimen di Media Online menggunakan metode Naive Bayes. *D'CARTESIAN: Jurnal ...*, September, 18–22. <https://ejournal.unsrat.ac.id/index.php/decartesian/article/view/36624%0Ahttps://ejournal.unsrat.ac.id/index.php/decartesian/article/viewFile/36624/39057>
- Sihombing, D. O. (2022). Implementasi Natural Language Processing (NLP) dan Algoritma Cosine Similarity dalam Penilaian Ujian Esai Otomatis. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 4(2), 396. <https://doi.org/10.30865/json.v4i2.5374>
- Sudianto, S., Sripamuji, A. D., Ramadhanti, I., Amalia, R. R., Saputra, J., & Prihatnowo, B. (2022). Penerapan Algoritma Support Vector Machine dan Multi-Layer Perceptron pada Klasisifikasi Topik Berita. *Jurnal Nasional Pendidikan Teknik Informatika : JANAPATI*, 11(2), 84–91. <https://ejournal.undiksha.ac.id/index.php/janapati/article/view/44151>
- Yudi, E., & Aditya, M. (2020). *Jurnal Nasional Teknologi dan Sistem Informasi Klasifikasi Dokumen Berita Menggunakan Algoritma Enhanced Confix Stripping Stemmer dan Naïve Bayes Classifier*. 02, 90–99.
- Zhafira, D. F., Rahayudi, B., & Indriati, I. (2021). 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*, 2(1), 55–63. <https://doi.org/10.25126/justsi.v2i1.24>

LAMPIRAN

Lampiran 1 Daftar *Stopword Removal*

Daftar <i>Stopword Removal</i>			
ada	bahwasanya	berkehendak	cukuplah
adalah	baik	berkeinginan	cuma
adanya	bakal	berkenaan	dahulu
adapun	bakalan	berlainan	dalam
agak	balik	berlalu	dan
agaknya	banyak	berlangsung	dapat
agar	bapak	berlebihan	dari
akan	baru	bermacam	daripada
akankah	bawah	bermacam-macam	datang
akhir	beberapa	bermaksud	dekat
akhiri	begini	bermula	demi
akhirnya	beginian	bersama	demikian
aku	beginikah	bersama-sama	demikianlah
akulah	beginilah	bersiap	dengan
amat	begitu	bersiap-siap	depan
amatlah	begitukah	bertanya	di
anda	begitulah	bertanya-tanya	dia
andalah	begitupun	berturut	diakhiri
antar	bekerja	berturut-turut	diakhirinya
antara	belakang	bertutur	dialah
antaranya	belakangan	berujar	diantara
apa	belum	berupa	diantaranya
apaan	belumah	besar	diberi
apabila	benar	betul	diberikan
apakah	benarkah	betulkah	diberikannya
apalagi	benarlah	biasa	dibuat
apatah	berada	biasanya	dibuatnya
artinya	berakhir	bila	didapat
asal	berakhirilah	bilakah	didatangkan
asalkan	berakhirnya	bisa	digunakan
atas	berapa	bisakah	diibaratkan
atau	berapakah	boleh	diibaratkannya
ataukah	berapalah	bolehkah	diingat
ataupun	berapapun	bolehlah	diingatkan
awal	berarti	buat	diinginkan
awalnya	berawal	bukan	dijawab
bagai	berbagai	bukankah	dijelaskan
bagaikan	berdatangan	bukanlah	dijelaskannya
bagaimana	beri	bukannya	dikarenakan
bagaimanakah	berikan	bulan	dikatakan
bagaimanapun	berikut	bung	dikatakannya
bagi	berikutnya	cara	dikerjakan
bagian	berjumlah	caranya	diketahui
bahkan	berkali-kali	cukup	diketuainya
bahwa	berkata	cukupkah	dikira
dilakukan	dulu	jikalau	kira-kira
dilalui	empat	juga	kiranya
dilihat	enggak	jumlah	kita
dimaksud	enggaknya	jumlahnya	kitalah

Lanjutan Lampiran 1

dimaksudkan	entah	justru	kok
dimaksudkannya	entahlah	kala	kurang
dimaksudnya	guna	kalau	lagi
diminta	gunakan	kalaulah	lagian
dimintai	hal	kalaupun	lah
dimisalkan	hampir	kalian	lain
dimulai	hanya	kami	lainnya
dimulailah	hanyalah	kamilah	lalu
dimulainya	hari	kamu	lama
dimungkinkan	harus	kamulah	lamanya
dini	haruslah	kan	lanjut
dipastikan	harusnya	kapan	lanjutnya
diperbuat	hendak	kapankah	lebih
diperbuatnya	hendaklah	kepun	lewat
dipergunakan	hendaknya	karena	lima
diperkirakan	hingga	karenanya	luar
diperlihatkan	ia	kasus	macam
diperlukan	ialah	kata	maka
diperlukannya	ibarat	katakan	makanya
dipersoalkan	ibaratkan	katakanlah	makin
dipertanyakan	ibaratnya	katanya	malah
dipunyai	ibu	ke	malahan
diri	ikut	keadaan	mampu
dirinya	ingat	kebetulan	mampukah
disampaikan	ingat-ingat	kecil	mana
disebut	ingin	kedua	manakala
disebutkan	inginkah	keduanya	manalagi
disebutkannya	inginkan	keinginan	masa
disini	ini	kelamaan	masalah
disinilah	inikah	kelihatan	masalahnya
ditambahkan	inilah	kelihatannya	masih
ditandaskan	itu	kelima	masihkah
ditanya	itukah	keluar	masing
ditanyai	itulah	kembali	masing-masing
ditanyakan	jadi	kemudian	mau
ditegaskan	jadilah	kemungkinan	maupun
ditujukan	jadinya	kemungkinannya	melainkan
ditunjuk	jangan	kenapa	melakukan
ditunjuki	jangan	kepada	melalui
ditunjukkan	janganlah	kepadanya	melihat
ditunjukkannya	jauh	kesampaian	melihatnya
ditunjuknya	jawab	keseluruhan	memang
dituturkan	jawaban	keseluruhannya	memastikan
dituturkannya	jawabnya	keterlalu	memberi
diucapkan	jas	ketika	memberikan
diucapkannya	jas	khususnya	membuat
diungkapkan	jas	kini	memerlukan
dong	jas	kinilah	memihak
dua	jika	kira	meminta
memintakan	menurut	pernah	sebisanya
memisalkan	menuturkan	persoalan	sebuah
memperbuat	menyampaikan	pertama	sebut

Lanjutan Lampiran 1

mempergunakan	menyangkut	pertama-tama	sebutlah
memperkirakan	menyatakan	pertanyaan	sebutnya
memperlihatkan	menyebutkan	pertanyakan	secara
mempersiapkan	menyeluruh	pihak	secukupnya
mempersoalkan	menyiapkan	pihaknya	sedang
mempertanyakan	merasa	pukul	sedangkan
mempunyai	mereka	pula	sedemikian
memulai	mereka	pun	sedikit
memungkinkan	merupakan	punya	sedikitnya
menaiki	meski	rasa	seenaknya
menambahkan	meskipun	rasanya	segala
menandaskan	meyakini	rata	segalanya
menanti	meyakinkan	rupanya	segera
menanti-nanti	minta	saat	seharusnya
menantikan	mirip	saatnya	sehingga
menanya	misal	saja	seingat
menanyai	misalkan	sajalah	sejak
menanyakan	misalnya	saling	sejauh
mendapat	mula	sama	sejenak
mendapatkan	mulai	sama-sama	sejumlah
mendatang	mulailah	sambil	sekadar
mendatangi	mulanya	sampai	sekadarnya
mendatangkan	mungkin	sampai-sampai	sekali
menegaskan	mungkinkah	sampaikan	sekali-kali
mengakhiri	nah	sana	sekalian
mengapa	naik	sangat	sekaligus
mengatakan	namun	sangatlah	sekalipun
mengatakannya	nanti	satu	sekarang
mengenai	nantinya	saya	sekarang
mengerjakan	nyaris	sayalah	sekecil
mengetahui	nyatanya	se	seketika
menggunakan	oleh	sebab	sekiranya
menghendaki	olehnya	sebabnya	sekitar
mengibaratkan	pada	sebagai	sekitarnya
mengibaratkannya	padahal	sebagaimana	sekarang-kurangnya
mengingat	padanya	sebagainya	sekarangnya
mengingatkan	pak	sebagian	sela
menginginkan	paling	sebaik	selain
mengira	panjang	sebaik-baiknya	selaku
mengucapkan	pantas	sebaiknya	selalu
mengucapkannya	para	sebaliknya	selama
mengungkapkan	pasti	sebanyak	selama-lamanya
menjadi	pastilah	sebegini	selamanya
menjawab	penting	sebegitu	selanjutnya
menjelaskan	pentingnya	sebelum	seluruh
menuju	per	sebelumnya	seluruhnya
menunjuk	percuma	sebenarnya	semacam
menunjuki	perlu	seberapa	semakin
menunjukkan	perlu	sebesar	semampu
menunjuknya	perlunya	sebetulnya	semampunya
semasa	setidak-tidaknya	terbanyak	ungkapnya
semasih	setidaknya	terdahulu	untuk

Lanjutan Lampiran 1

semata	setinggi	terdapat	usah
semata-mata	seusai	terdiri	usai
semaunya	sewaktu	terhadap	waduh
sementara	siap	terhadapnya	wah
semisal	siapa	teringat	wahai
semisalnya	siapakah	teringat-ingat	waktu
sempat	siapapun	terjadi	waktunya
semua	sini	terjadilah	walau
semuanya	sinilah	terjadinya	walaupun
semula	soal	terkira	wong
sendiri	soalnya	terlalu	yaitu
sendirian	suatu	terlebih	yakin
sendirinya	sudah	terlihat	yakni
seolah	sudahkah	termasuk	yang
seolah-olah	sudahlah	ternyata	polisi
seorang	supaya	tersampaikan	kasus
sepanjang	tadi	tersebut	kejahatan
sepantasnya	tadinya	tersebutlah	tindak
sepantasnyalah	tahu	tertentu	pelaku
seperlunya	tahun	tertuju	lokasi
seperti	tak	terus	warga
sepertinya	tambah	terutama	kantor
sepihak	tambahnya	tetap	menurut
sering	tampak	tetapi	ditemukan
seringnya	tampaknya	tiap	masyarakat
serta	tandas	tiba	melakukan
serupa	tandasnya	tiba-tiba	terjadi
sesaat	tanpa	tidak	petugas
sesama	tanya	tidakkah	korban
sesampai	tanyakan	tidaklah	menyatakan
sesegera	tanyanya	tiga	menyebutkan
sesekali	tapi	tinggi	terlibat
seseorang	tegas	toh	melibatkan
sesuatu	tegasnya	tunjuk	diduga
sesuatunya	telah	turut	mengatakan
sesudah	tempat	tutur	memiliki
sesudahnya	tengah	tuturnya	mengaku
setelah	tentang	ucap	mengungkapkan
setempat	tentu	ucapnya	penyidikan
setengah	tentulah	ujar	penangkapan
seterusnya	tentunya	ujarnya	mengakui
setiap	tepat	umum	menyelidiki
setiba	terakhir	umumnya	mencurigai
setibanya	terasa	ungkap	dilaporkan

Lampiran 2 Daftar Data Artikel Berita

No	Judul Artikel Berita	Kategori
1	3 Perampok Bersenpi Sandera Pangusaha di Muba Ditangkap, 1 Tewas Ditembak	Perampokan
2	Keroyok dan Rampok Korban, 7 Anggota Geng Motor di Banyuwangi Diringkus	Perampokan
3	Anjing di Surabaya Viral Mati Usai Lawan Perampok, Ternyata...	Perampokan
4	Kronologi Sopir Truk 'Dibuang' di Bogor Rekayasa Penggelapan 25 Ton Gula	Perampokan
5	Otak Perampokan Pengusaha Sawit di Muba Tewas Ditembak, Begini Perannya	Perampokan
6	Ketahuhan Sembunyi di TPU, Seorang Tersangka Perampokan Sempat Diamuk Massa	Perampokan
7	8 Perampok Bersenjata di Serang Dilumpuhkan, Terakhir Jarah Rp 200 Juta dan Emas 85 Gram Milik Juragan Beras	Perampokan
8	Polda Metro Tangkap 5 Pelaku Lain Kasus Rekayasa Perampokan Truk Gula di Tol Cikupa	Perampokan
9	Perampokan Sopir Truk yang Dibuang di Bogor, Ternyata Rekayasa	Perampokan
10	Bobol Brankas Senilai Rp 90 Juta, Komplotan Perampok Ditangkap, 2 di Antaranya Pecatan TNI	Perampokan
11	Penjual Tompo di Malang Jadi Korban Perampokan, Uang Rp 19 Juta Hasil Menabung Lenyap	Perampokan
12	Kasus Dugaan Perampokan dan Pembunuhan di Mamasa, Polisi Bentuk Tim Khusus hingga TKP Dijaga Ketat	Perampokan
13	Polisi Tak Temukan Indikasi Perampokan dalam Kasus Penembakan Bank Swasta di Cengkareng	Perampokan
14	Ze Valente, Aubameyang, dan Deretan Pesepak Bola yang Jadi Korban Perampokan	Perampokan
15	Cerita Kakek Penjual Tompo Jadi Korban Perampokan, Tak Curiga Dagangan Diborong hingga Uang Rp 19 Juta Tabungan Lenyap	Perampokan
16	Terkuak, Pecatan TNI Pasok Senpi Pabrikasi ke Perampok Toko Emas Serpong	Perampokan
17	Rangkaian Penangkapan 4 Perampok Toko Emas, Dibekuk di Bogor hingga Jateng	Perampokan
18	Ini Tampang 4 Tersangka Perampokan Toko Emas di Tangerang Raya	Perampokan
19	Aksi Nekat Perampok Toko Emas di Siang Bolong Kini Diburu Polisi	Perampokan
20	Rentetan 3 Aksi yang Dilakukan Kelompok Perampok Toko Emas di Serpong	Perampokan
21	Perampok Alfamart Teuku Umar Denpasar Dibekuk, Kedua Kakinya Ditembak	Perampokan
22	Detik-detik Perampok Bersenpi Beraksi di Jalinsum Sumsel Gasak Rp 350 juta	Perampokan
23	Pria di Bogor Ngaku Korban Rampok Modus Kempis Ban, Duit Rp 400 Juta Raib	Perampokan
24	Polisi Buru Dua Pelaku Perampokan di Desa Pegayaman	Perampokan
25	Perampok yang Buron 6 Bulan di Banyuwangi Keok di Tangan Polisi	Perampokan
26	Perampok Berpistol di SPBU Sidoarjo Pakai Motor Matic-Berhelm Merah	Perampokan
27	Polisi Endus Kejanggalan Aksi Perampokan di SPBU Lubuk Batang Sumsel	Perampokan
28	Perampok di Agam Bawa Kabur Emas Sebanyak 2 Kg	Perampokan
29	Kronologi Perampokan di Cikokol, Tepergok Tuan Rumah dan Sempat Berkelahi	Perampokan
30	Gagalkan Aksi Perampok Bersenjata, Pemilik Rumah di Cikokol Tangerang Alami Luka Robek	Perampokan
31	3 Bulan Buron, 4 Perampok Uang Rp 320 Juta Milik Pesantren Mudi Samalanga Aceh Ditangkap	Perampokan
32	Kronologi Perampokan Minimarket di Leuwiliang Bogor, Pelaku Bermasker dan Todongkan Pedang	Perampokan
33	Perampok BPR Kota Kediri Ditangkap, Motifnya karena Terjerat Utang dan Judi Online	Perampokan
34	Perampok Bersenjata Api Sasar Minimarket 24 Jam di Pekanbaru, Sudah 2 Tempat Jadi Korban	Perampokan
35	Perampok Sadis di Jambi Ternyata Tetangga Sendiri, Ditangkap 2 Jam Setelah Kejadian	Perampokan
36	Detik-detik Warga di Banyumas Kejar dan Tangkap Perampok Uang Rp 70 Juta	Perampokan
37	Polisi Pastikan Belum Terjadi Perampokan dalam Kasus Penembakan Hansip di Tambora	Perampokan

Lanjutan Lampiran 2

38	Minimarket Pagedangan Tangerang Dirampok, Komplotan Tunggu Warga Asyik Nonton Piala Dunia	Perampokan
39	Minimarket di Kabupaten Bandung Dirampok, Puluhan Juta Rupiah Hilang	Perampokan
40	Polisi Diserang Saat Tangkap Perampok di Sumsel, Mobil Dirusak dan Pelaku Kabur	Perampokan
41	Polisi Tangkap Perampok Toko Perhiasan di Jember, Sita 1,5 Kilogram Emas	Perampokan
42	3 Perampok yang Bacok Korbannya di Riau Ditangkap, Satu Pelaku Dilumpuhkan Polisi	Perampokan
43	Perampok Rumah Dinas Wali Kota Blitar Rusak CCTV dan Bawa Dekoder	Perampokan
44	Tak Hanya Disekap, Wali Kota Blitar Juga Sempat Dianiaya Perampok	Perampokan
45	Polisi: Perampok Rumah Dinas Wali Kota Blitar Pakai Mobil Pelat Merah	Perampokan
46	Perampokan Rumah Dinas Wali Kota Blitar, Warga Dengar Teriakan Minta Tolong	Perampokan
47	Wali Kota Blitar dan Istri Disekap Perampok di Rumah Dinas, Uang Rp 400 Juta dan Perhiasan Raib	Perampokan
48	Rumah Dinas Dirampok, Wali Kota Blitar dan Istri Sempat Disekap	Perampokan
49	Pria Tangerang Rampok 2 Minimarket di Madiun Bermodal Pistol Korek Api	Perampokan
50	Mantan Walkot Rampok Walkot Blitar, Rencana Digodok dari dalam Penjara	Perampokan
51	Komplotan Pencuri Gasak Uang-Emas Total Rp 1 M Diringkus di Semarang	Perampokan
52	Kisah Polisi Bekuk Bule Rusia Komplotan Perampok di Bali Dijadikan Film	Perampokan
53	Pakai Parang, 3 Pria Rampok Motor Pasutri Saat Hendak ke Pasar di Medan	Perampokan
54	3 Perampok Motor Pasutri di Medan Ditangkap, Begini Perannya	Perampokan
55	Dua Cewek Rampok Minimarket di Nganjuk, Bawa Celurit Gasak Rp 47 juta	Perampokan
56	4 Perampok di Makassar Todong Kasir Minimarket Pakai Samurai Ditangkap	Perampokan
57	Perampokan Sadis di Kedungreja Cilacap: Kronologi dan Jumlah Korban	Perampokan
58	Perampok yang Buang Anggota TNI AL Ternyata Sudah Beraksi di 10 TKP	Perampokan
59	Perampok Pakai Duit Pengusaha Sukabumi untuk Bayar Utang-Beli Motor	Perampokan
60	Polisi Tangkap 3 Perampok di Rest Area Tol Ngawi, Satu Pelaku Ditembak	Perampokan
61	Kekasih Ungkap Brigadir J Terima Ancaman Pembunuhan Seminggu Sebelum Tewas	Pembunuhan
62	Pria Diduga Korban Pengeroyokan Ditemukan Tewas di Sidoarjo	Pembunuhan
63	Adik-Kakak Jadi Tersangka Pembunuhan gegara Percikan Ludah	Pembunuhan
64	Pengacara Ungkap Brigadir J Dapat Ancaman Pembunuhan Sejak Juni dan Juli	Pembunuhan
65	Motif Pembunuhan Perempuan di Kamar Mandi Hotel di Surabaya, Pelaku Tergiur Uang Rp 20 Juta	Pembunuhan
66	Sadisnya Aksi Pembunuhan oleh Sejoli di Karang Tengah, Sayat Wajah Korban Pakai Pisau agar Tak Dikenali	Pembunuhan
67	Kasus Pembunuhan Ibu dan Anak di Subang, Polisi Amankan Seorang Berinisial S, Siapa Dia?	Pembunuhan
68	Duga Brigadir J Jadi Korban Pembunuhan Berencana, Keluarga Laporkan ke Bareskrim Hari Ini	Pembunuhan
69	Pelaku Pembunuhan Petugas Kebersihan Palembang Ditangkap	Pembunuhan
70	Kasus Kematian Brigadir J, Dugaan Pembunuhan Berencana dan Rencana Ekshumasi Jenazah	Pembunuhan
71	Kasus Pembunuhan Ibu dan Anak di Subang Kembali Diusut, Harapan Baru bagi Keluarga	Pembunuhan
72	Pelaku Pembunuhan Pelajar SMP di Magelang Teman Korban, Ini Motifnya	Pembunuhan
73	Kasus Pembunuhan Ibu dan Anak di Subang, Seorang ABK Diperiksa	Pembunuhan
74	Setahun Kasus Pembunuhan Ibu dan Anak di Subang, Seratusan Saksi Diperiksa, tapi Pelaku Tak Kunjung Tertangkap	Pembunuhan
75	3 Kasus Pembunuhan Misterius yang Belum Terungkap	Pembunuhan
76	Kasus Pembunuhan Ibu dan Anak di Subang Hampir Setahun, Kapolda Jabar: Kami Masih Berupaya, Butuh Pembuktian Mendalam	Pembunuhan
77	Wanita di Magelang Jadi Korban Pembunuhan, Pelaku Ditangkap di Ngawi	Pembunuhan
78	Penjelasan Polisi Usai Disebut "Cueki" Laporan KDRT Korban Pembunuhan Suami di Bekasi	Pembunuhan
79	Polisi Lepas ABK yang Ada Dalam TKP Pembunuhan Ibu dan Anak di Subang	Pembunuhan

Lanjutan Lampiran 2

80	TNI AD Periksa 6 Prajurit yang Diduga Terlibat Kasus Pembunuhan Sadis di Mimika	Pembunuhan
81	4 Warga Jadi Korban Pembunuhan Sadis di Mimika Papua, 3 Pelaku Ditangkap!	Pembunuhan
82	Detik-detik Polisi Amankan Pria Kaitan Pembunuhan Ibu-Anak Subang	Pembunuhan
83	Pembunuh Remaja di Pasar Simpang Ditangkap!	Pembunuhan
84	Horor Aksi 'Dukun' di Jabar Bunuh Warga: Pakai Nisan-Racun Sianida	Pembunuhan
85	Pembunuhan PNS Saksi Korupsi Kian Terang, Polisi Minta Pelaku Serahkan Diri	Pembunuhan
86	Jabar X-Files: Selimut Misteri 3 Kasus Pembunuhan Sadis Para Wanita	Pembunuhan
87	Pelaku-Motif Pembunuhan Sadis Pasutri di Kalimantan Tengah Masih Teka Teki	Pembunuhan
88	Pembunuh Marbut Masjid di Indramayu Ditangkap!	Pembunuhan
89	Polisi Ungkap Detik-detik Aksi Sadis Suami Bunuh Istri di Pemalang	Pembunuhan
90	Pelaku Pembunuhan Wanita Dalam Tas Merah di Gresik Masih Gelap	Pembunuhan
91	Mayat Pria Tangan Terikat-Kepala Ditutup Karung Diduga Korban Pembunuhan	Pembunuhan
92	Tega Bunuh Wanita-Mayatnya Dibungkus Plastik karena Pelaku Sakit Hati	Pembunuhan
93	Hari Anti Hukuman Mati Sedunia 10 Oktober 2022, Bagaimana di Indonesia?	Pembunuhan
94	Kejadian Bunuh Diri di Surabaya Capai 11 Kasus hingga Oktober 2022	Pembunuhan
95	Kronologi Pembunuhan Wanita dalam Karung di Bangsri Jepara gegara Utang	Pembunuhan
96	Senyum Pembunuh Saat Dorong Troli Berisi Jasad Wanita Terekam CCTV	Pembunuhan
97	Korban Pembunuhan di Soreang Ternyata Mahasiswa Unpad, Warga Garut	Pembunuhan
98	5 Fakta Pembunuhan Mahasiswa Unpad di Bandung	Pembunuhan
99	Kronologi Pembunuhan Kakek 78 tahun di Yogyakarta, Dijerat Tali oleh Cucu Dalam Mobil di Parkiran Restoran	Pembunuhan
100	Anak Kedua Terduga Pembunuh Satu Keluarga di Magelang Beli Racun dari Online, Dicampur dalam Teh dan Kopi Korban	Pembunuhan
101	Kasus Pembunuhan Mahasiswa di Palembang, Mayat Disimpan di Bagasi Mobil Semalaman Sebelum Dibakar	Pembunuhan
102	Kronologi Pembunuhan Sadis Adik Ipar di Indramayu, Dibacok dari Belakang	Pembunuhan
103	Pembunuhan Sadis Ibu Hamil di Yogyakarta, Dilempar dalam Kondisi Hidup dari Tebing Pantai karena Tolak Aborsi	Pembunuhan
104	Misteri Mayat Perempuan Dalam Karung di Bogor, Dibunuh Selingkuhan yang Juga Guru Ngaji Anak Korban	Pembunuhan
105	Seorang Wanita Ditemukan Tewas Bersimbah Darah di Madiun, Diduga Korban Pembunuhan	Pembunuhan
106	Bunuh Kepala Total Buah Serpong, Pelaku Pura-pura Tidak Tahu Ada Pembunuhan	Pembunuhan
107	WNA Pembunuh Perempuan Bertato Kupu-kupu Terancam Hukuman Seumur Hidup	Pembunuhan
108	Petarung MMA Bunuh Kakak Sendiri gegara Masalah Kompromi	Pembunuhan
109	2 Orang Pelaku Pembunuhan Pria Terikat Leher di Tangerang Ditangkap!	Pembunuhan
110	Fakta-fakta Pembunuh Sekeluarga di Bekasi: Serial Killer Supranatural	Pembunuhan
111	3 Kabar Terkini Sekeluarga Keracunan di Bekasi Diduga Korban Pembunuhan	Pembunuhan
112	Jabar Hari Ini: Akhir Kisah Pembunuhan Nanay Berlyn	Pembunuhan
113	Kabar Terkini Duo Pembunuh Bos Ayam Goreng Ternyata Juga Mencuri	Pembunuhan
114	Pembunuhan oleh Pacar: Kejanggalan Status Wa-Ayah Tak Tahu Anaknya Hamil	Pembunuhan
115	Kasus Pembunuhan Dokter Spesialis di Nabire, Jasad Korban Diperkosa Pelaku	Pembunuhan
116	Bunuh Pria di Kaltim, Jordi Baru 21 Hari Bebas dari Tahanan karena RJ	Pembunuhan
117	Nyawa Rani Melayang Ditusuk Pedang Suami di Hari Lebaran	Pembunuhan
118	Kakak Adik Bunuh Wanita Dalam Karung di Jakut Dijerat Pembunuhan Berencana	Pembunuhan
119	Pembunuh Ambil Duit Rp 15 Ribu Usai Habisi Ibu Anggota DPR RI	Pembunuhan
120	Rudolf Pembunuh 'Tersenyum di Lift' Jalani Sidang Putusan Sela Hari Ini	Pembunuhan
121	Polda Sulteng Sebut Kasus Perkosaan Anak 15 Tahun oleh 11 Pria dengan Persetujuan, Ini Kata Ahli	pemeriksaan
122	Update Kasus Pemeriksaan Anak di Parimo, Pelaku yang Buron Ditangkap di Kendari	pemeriksaan
123	Kronologi Pemeriksaan dan Perampokan SPG Mobil di Cibubur	pemeriksaan
124	Saat Oknum Perwira Polisi Jadi Tersangka Pemeriksaan Anak 15 Tahun di Sulteng...	pemeriksaan

Lanjutan Lampiran 2

125	Lemahnya Penegakan Hukum Kasus Pemerksaan Anak di Cipayung, Ibu Korban Malah Dimarahi Polisi	pemerksaan
126	Polisi Selidiki Kasus Pelajar Banyuwangi Diperksa 3 Orang hingga Melahirkan	pemerksaan
127	Fakta Lain! Korban Sempat Diperksa Berulang dan Diancam Pelaku	pemerksaan
128	Kronologi Pernikahan Pelajar Banyuwangi Korban Perksaan Diduga 'Settingan'	pemerksaan
129	Pengakuan yang Menyayat Hati dari Para Korban Pemerksaan Bechi	pemerksaan
130	Siswi SMP di Sragen 17 Kali Diperksa Ayah Tiri hingga Melahirkan, Terungkap dari Tes DNA	pemerksaan
131	Coba Perksa Seorang Remaja, Pria di Rote Ndao Terancam 5 Tahun Penjara	pemerksaan
132	9 Pelaku Pemerksaan Siswi SMP di NTT Ditangkap, 1 Orang Masih Kabur	pemerksaan
133	4 Pemuda di Alor Diduga Perksa Bergilir Siswi SMA di Kompleks PAUD, Korban Sempat Dicekoki Miras	pemerksaan
134	Residivis Pencabulan Anak Sekap dan Perksa Siswi SMP di Pati hingga Hamil, Ditangkap Saat Kabur ke Papua	pemerksaan
135	Pelaku yang Sekap dan Perksa Siswi SMP di Pati hingga Hamil Ditangkap, Ternyata Residivis Pencabulan Anak di Bawah Umur	pemerksaan
136	Dugaan hingga Kejanggalan Isu Pemerksaan Istri Sambo oleh Yoshua	pemerksaan
137	Fakta Baru Pemerksa ABG di Buleleng Ternyata Seorang Duda	pemerksaan
138	Kemensos Fasilitasi Kebutuhan Korban Pemerksaan Ayah Tiri di Kota Batu	pemerksaan
139	Pengacara Sekolah: Kasus Pemerksaan Ini Heboh gegara Hotman Paris	pemerksaan
140	Pihak Sekolah Sebut Kasus Siswi SD Diperksa Kepsek Diskenario Ibu Korban	pemerksaan
141	Kepala Lingkungan di Luwu Dilaporkan Perksa Bocah 11 Tahun Berkali-kali, Korban Diimingi Uang Rp 150.000	pemerksaan
142	Iming-imingi Uang Rp 5.000, Pria di Buleleng Perksa Bocah SD	pemerksaan
143	Kasus Pemerksaan Bocah 8 Tahun di Ambon, Pelaku dan Korban Tetangga Dekat	pemerksaan
144	Sambil Menangis Histeris, Korban Pemerksaan dan Keluarga Giring Terduga Pelaku ke Mapolres Depok	pemerksaan
145	Marak Kasus Pemerksaan di Jakut, Pemkot Didesak Bentuk Sparta	pemerksaan
146	Kasus Pemerksaan Bocah SD di Ciputat, Pelaku Salurkan Hawa Nafsu Setelah Bercerai	pemerksaan
147	Seorang Wanita Banyuwangi Jadi Korban Pemerksaan Tamu di Rumahnya	pemerksaan
148	Alasan Polisi Belum Bisa Tangkap 9 Pemerksa Anak 13 Tahun di Sampang	pemerksaan
149	Diajak Jalan-jalan, Remaja di Dompur Jadi Korban Percobaan Pemerksaan	pemerksaan
150	Kasus Pemerksaan, Dua PNS Kemenkop UKM Dipecat!	pemerksaan
151	Polisi Akan Cek Dugaan Obstruction of Justice Kasus Perksaan di Kemenkop	pemerksaan
152	Biadab Pria Kaltim Perksa Putrinya 4 Kali dalam Sepekan hingga Korban Hamil	pemerksaan
153	Komnas Perempuan Minta TNI Terus Dampingi Prajurit Kostrad Korban Pemerksaan Perwira Paspampres	pemerksaan
154	Pangkostrad Angkat Bicara Terkait Prajuritnya yang Diduga Diperksa Perwira Paspampres	pemerksaan
155	Kaleidoskop 2022: Perksa 13 Santri, Herry Wirawan Divonis Hukuman Mati	pemerksaan
156	Kabareskrim Siap Ambil Alih Kasus Pemerksaan Pegawai Kemenkop Jika Mandek	pemerksaan
157	Isyarat Damai Kasus Pemerksaan ABG 15 Tahun di Lombok Barat	pemerksaan
158	Pria Sumsel 6 Kali Perksa Anak 12 Tahun, Terungkap Usai Dijebak Ibu Korban	pemerksaan
159	Ini Tampang Pemerksa Santriwati di Masjid Deli Serdang	pemerksaan
160	Kala Ibu di Palembang Menyamar untuk Ungkap Pemerksaan Terhadap Anaknya	pemerksaan
161	Miris! Perksaan Anak oleh 6 Pria di Brebes Berakhir Mediasi di Rumah Kades	pemerksaan
162	Polisi Ajukan Restitusi untuk Korban Pemerksaan di Pinggir Tol Jakarta-Merak	pemerksaan
163	Dalam Masa Pemulihan, Santriwati Korban Pemerksaan Ustadz di Depok Masih Takut Bertemu Orang	pemerksaan
164	Wanita Bos Rental PS Diduga Cabuli 17 Anak Mayoritas Laki-laki, Pelaku Sempat Mengaku Jadi Korban	pemerksaan

Lanjutan Lampiran 2

165	Siswi SMP di Bone Tewas Diperkosa Ramai-ramai, Polisi Tangkap Teman Sekolah Korban	pemerksosaan
166	6 Pelaku Pemerksosaan Siswi SMP di Konawe Ditangkap Polisi, Satu Pelaku Buron	pemerksosaan
167	Polisi Selidiki Pemerksosaan Perempuan di Semak-semak Tol Jakarta-Tangerang	pemerksosaan
168	Terkuak Alasan Polisi Pulangkan Remaja Pemerksosa Siswi SMP Bone hingga Tewas	pemerksosaan
169	Laporan Yunita Tersangka Cabul 17 Anak di Jambi soal Pemerksosaan Dihentikan	pemerksosaan
170	Pria Bima Dibekuk gegara Hendak Perkosa IRT di Kebun Jagung	pemerksosaan
171	Nestapa Dosen Jadi Korban Penipuan-Pemerksosaan di Surabaya	pemerksosaan
172	Bejat! Dalih Antar Tagih Utang Mantan, 3 Pemuda Pekalongan Perkosa Gadis ABG	pemerksosaan
173	Akal-akalan Yunita Agar Disebut Jadi Korban Pemerksosaan	pemerksosaan
174	Kasus Pemerksosaan ABG Dilimpahkan ke Kejari, Tersangka Masih Mengelak	pemerksosaan
175	Detik-detik Pria Pemerksosa ABG di Jeneponto Ditangkap	pemerksosaan
176	Satu Pemerksosa Siswi SMP di Surabaya hingga Hamil 5 Bulan Ditangkap	pemerksosaan
177	Polisi Tangkap 1 dari 10 Tersangka Pemerksosa 2 Siswi di Asahan	pemerksosaan
178	Fransisca, Gadis Cilik Korban Pemerksosaan Mei 1998 dan Cerita yang Kian Terkubur	pemerksosaan
179	Terbongkarnya Kasus Pemerksosaan Gadis 16 Tahun oleh 11 Pria di Sulteng	pemerksosaan
180	Kronologi Gadis 16 Tahun di Sulteng Diperkosa 11 Orang, Ini Pelakunya	pemerksosaan
181	Tersangka Baru Kasus Garuda Indonesia: Sosok, Peran, Duduk Perkara	Korupsi
182	Negara Rugi Rp 10 M di Kasus Penggelapan Pajak Samsat Kelapa Dua	Korupsi
183	Tersangka Korupsi Retribusi Taman Kartini, Oknum ASN Pemkab Rembang Ditahan	Korupsi
184	PT Titan Dilaporkan ke Bareskrim Terkait Kasus Dugaan Kredit Macet	Korupsi
185	Andre Rosiade Apresiasi Gercep Erick-Jaksa Agung Tangani Kasus Garuda	Korupsi
186	Kadis Kesehatan Padang Sidempuan Tersangka Korupsi Dana COVID-19	Korupsi
187	Kejagung Periksa 9 Saksi dalam Kasus Korupsi Izin Ekspor Minyak Goreng	Korupsi
188	Tersangka Korupsi Kabur, KPK Bantah Ada Kebocoran Informasi di Internal	Korupsi
189	KPK Tetapkan 3 Orang Tersangka Korupsi Pembangunan Stadion Mandala Krida	Korupsi
190	Kejagung Tetapkan 5 Tersangka Korupsi Pembangunan Pabrik PT Krakatau Steel, Salah Satunya Eks Dirut	Korupsi
191	Kejagung Tetapkan 4 Tersangka Kasus Dugaan Korupsi Waskita Beton	Korupsi
192	Kejati Bali Geledah Rumah Tersangka Korupsi LPD Sangeh Badung	Korupsi
193	KPK Sedang Usut Kasus Besar Pertambangan, Nama Karen Agustiawan Disebut	Korupsi
194	Kades Roomo Gresik Jadi Tersangka Korupsi ADD Rp 270 Juta	Korupsi
195	Bupati Tulungagung Maryoto Birowo Kembali Diperiksa KPK	Korupsi
196	13 Terdakwa Korupsi Ganti Rugi Lahan Tol Padang-Pekanbaru Divonis Bebas	Korupsi
197	Hari Ini, Surya Darmadi Jalani Sidang Perdana Kasus Korupsi Rp 104 Triliun	Korupsi
198	Deretan Skandal Lukas Enembe, dari Dugaan Korupsi hingga Judi	Korupsi
199	Kejagung Tetapkan 5 Tersangka Kasus Korupsi Pembelian Tanah PT Adhi Persada Realti	Korupsi
200	23 Koruptor Bebas Bersyarat pada 6 September, Ada Pinangki dan Patrialis Akbar	Korupsi
201	Bantu Suami Bisnis Narkoba dari Lapas, Wanita di Semarang Ditangkap	Korupsi
202	Dakwaan Korupsi Heli AW-101, Negara Rugi Rp 738 Miliar!	Korupsi
203	Nasib Eks Rektor UIN Suska Riau: Dipecat Menag, Kini Tersangka Korupsi	Korupsi
204	Disebut Terima Rp 17 M di Kasus AW-101, Eks KSAU Agus Supriatna: Ngarang	Korupsi
205	KPK Panggil Kepala BPK Bali soal Dugaan Suap Izin Infrastruktur di Sulsel	Korupsi
206	KPK Panggil Dirjen Dikti dan Rektor ITS di Kasus Suap Rektor Unila	Korupsi
207	Mardani Maming Didakwa Terima Rp 118 M Terkait Izin Tambang	Korupsi
208	Klaim Suap Rp 100 Juta Rektor Unila ke Muktamar Ditepis Nahdlatul Ulama	Korupsi
209	Saksi Curhat Beli Rumah Eks Pejabat Samsat Banten tapi Malah Disita	Korupsi
210	KPK: Dito Mahendra Mangkir Panggilan Sebagai Saksi Kasus Korupsi Nurhadi	Korupsi
211	Hukuman Koruptor Dipangkas dalam RKUHP, Korupsi Bukan Lagi "Extraordinary Crime"?	Korupsi
212	Hadiri Hakordia 2022, Wapres: Sama seperti Covid-19, Korupsi Musuh Semua Bangsa	Korupsi
213	Litbang "Kompas": Publik Tak Setuju Mantan Napi Korupsi Jadi Caleg	Korupsi

Lanjutan Lampiran 2

214	Kejagung Tangani 8 Kasus Besar Selama 2022: dari Korupsi Pengadaan Pesawat PT Garuda hingga Izin Ekspor CPO	Korupsi
215	Hakordia 2022: Ironi Adik-Kakak Terjerat Kasus Korupsi	Korupsi
216	2 Kali Tersangka dan Ditahan, Eks Bupati Inhil Ajukan Praperadilan	Korupsi
217	Wow! Anggaran Pembuatan Konten Masjid Al Jabbar Capai Rp 15 M	Korupsi
218	4 Terdakwa Korupsi Pajak Samsat Kelapa Dua Banten Dituntut 8 Tahun Bui	Korupsi
219	Hakim Tolak Praperadilan Eks Bupati Inhil Indra Mukhlis!	Korupsi
220	Korupsi Kaus Kaki, Eks Kasatpol PP Asnawi Gay Dihukum 2 Tahun Bui	Korupsi
221	Firli: Tempat Persembunyian DPO KPK Tak Terbatas di Wilayah Indonesia	Korupsi
222	John Irfan Kenway Dituntut 15 Tahun Bui di Kasus Heli AW-101	Korupsi
223	Eks Penyidik KPK Sebut Janji Jokowi Perkuat Pemberantasan Korupsi Hanya Basa-basi	Korupsi
224	Kejagung Sita Dokumen hingga Motor Ducati Terkait Kasus Korupsi BTS 4G Kominfo	Korupsi
225	Menkominfo dan Sang Adik dalam Pusaran Kasus Korupsi BTS 4G Kominfo	Korupsi
226	Kejagung Sita Rp 41 M Terkait Kasus Korupsi Waskita Karya	Korupsi
227	Jadi Tersangka Korupsi Bansos, Kuncoro Wibowo Pernah Dapat Penghargaan BUMN	Korupsi
228	KPK Tahan Tersangka Kasus Rintang Penyidikan Suap Eks Bupati Buru Selatan	Korupsi
229	Sunjaya Pasang Badan untuk GM Hyundai Muluskan Proyek PLTU 2 Cirebon	Korupsi
230	Dana Desa Bari Diduga 4 Tahun Dikorupsi Hampir Rp 500 Juta	Korupsi
231	MA Gandakan 6 Kali Lipat Hukuman Duo Koruptor Jalan dari Aceh Ini	Korupsi
232	Kebocoran Dokumen Korupsi di ESDM, Eks Komisioner Sebut KPK Dijegal	Korupsi
233	Korupsi Berjemaah di DJKA Kemenhub, Para Tersangka Diduga Terima Rp 14,5 M	Korupsi
234	Kilas Balik Kasus Anas Urbaningrum: Korupsi Proyek Hambalang, Hukuman Dipangkas, Kini Bebas	Korupsi
235	Selain Rudiantara, Empat Saksi Lainnya Hadir dalam Sidang Kasus Dugaan Korupsi Satelit Kemenhan	Korupsi
236	Dirut PT Waskita Karya Jadi Tersangka Korupsi, Ini Perannya	Korupsi
237	Dukung Pengusutan Kasus Korupsi, Pelindo Terus Benahi Pengelolaan Dapen	Korupsi
238	Kejagung Tetapkan Tersangka Baru Korupsi BTS 4G, Siapa?	Korupsi
239	Duduk Perkara Eks Kades di Sumsel Korupsi untuk Sewa Wanita Open BO	Korupsi
240	PNS Disdikpora Bantul Ditahan Kasus Korupsi Stadion Sultan Agung	Korupsi
241	DJ Joice Rehabilitasi Rawat Jalan di BNNK Jaksel	Narkoba
242	Pemuda Pengedar 92 Paket Sabu di Gilimanuk Ditangkap Polisi	Narkoba
243	Terkuak Sosok Pengusaha Busana yang Nyabu Bareng Pegawai Imigrasi Jember	Narkoba
244	Polisi Tangkap Bandar Narkoba di Madina, 7,8 Kg Ganja Disita	Narkoba
245	Duh, Oknum Pegawai Imigrasi Jember Ditangkap Nyabu Bareng Pengusaha	Narkoba
246	8 Pengedar Narkoba di Wilayah Bogor dan Depok Ditangkap, 2 Residivis	Narkoba
247	Pecatan Polisi di Jembrana Edarkan Narkoba, Bakar BB saat Ditangkap	Narkoba
248	Pengedar Sewa Rumah di Kampung Ambon untuk Distribusikan Narkoba, Kini Ditangkap Polisi	Narkoba
249	Bandar Narkoba Dibekuk, Simpan 10 Gram Sabu di Kantong Celana	Narkoba
250	Gelagatnya Mencurigakan, 2 Kurir Narkoba Ditangkap di Kampung Ambon, 81 Paket Sabu Disita Polisi	Narkoba
251	Baru Bebas dari Penjara, Residivis Kasus Narkoba di Kota Malang Kembali Ditangkap	Narkoba
252	Polisi Tangkap 5 Anggota Sindikat Narkoba, 44 Kg Ganja Disita	Narkoba
253	3 Pengedar Narkoba di Tegal Ditangkap Saat Bungkus Ganja Siap Edar	Narkoba
254	Bareskrim Bongkar Peredaran Narkoba Jaringan Internasional, Polisi Aktif Jadi Kurir	Narkoba
255	Sopir Truk Maut di Bekasi Negatif Narkoba, Habis Minum Obat Asam Urat	Narkoba
256	Eks Kapolres Bandara Soetta Kombes Edwin Hatorangan Diberhentikan Tidak Hormat!	Narkoba
257	7 Fakta Anak Imam S Arifin Gelapkan Motor Diduga demi Narkoba	Narkoba
258	Bandar Sabu Kelas Kakap Dibekuk bersama Istri dan 2 Sepupu di Mataram	Narkoba
259	Napi Kendalikan Peredaran Ekstasi dari Lapas, Begini Respons Kemenkumham	Narkoba
260	Terkuak! Istri Bandar di Mataram Berperan Jadi Marketing Sabu	Narkoba

Lanjutan Lampiran 2

261	Atasi Penyalahgunaan Narkoba, Walkot Samarinda Ajak Camat dan Lurah Perkuat Sinergi dalam Rakor Kotan	Narkoba
262	Kronologi Penangkapan Irjen Teddy Minahasa Diduga Terlibat Narkoba, Berawal dari Penyidikan Polda Metro Jaya	Narkoba
263	Bareskrim Amankan 270,283 Kg Sabu dari 4 Kasus Narkoba pada September-Oktober 2022	Narkoba
264	8 Pengedar Narkoba Ditangkap di Kampung Ambon, Berawal dari Laporan Orang Tak Bayar Makanan dan Rampas HP Ojol	Narkoba
265	Polisi Tangkap Dua Pengedar Narkoba yang Dikendalikan Napi dari Lapas di Bandung	Narkoba
266	Tergiur Upah, Pengantin Baru Nekad Jadi Kurir Narkoba Internasional	Narkoba
267	Bisa-bisanya Wihara Thailand Kosong Buntut Para Biksu Dites Narkoba	Narkoba
268	Penyelundupan Ganja dan Sabu ke LP Cipinang Via Truk Sampah Digagalkan	Narkoba
269	'Empuknya' Sukabumi Jadi Sasaran Bandar Narkoba	Narkoba
270	Nekat Edarkan Sabu, Kuli Bangunan di Cirebon Ditangkap Polisi	Narkoba
271	Polresta Barelang Gagalkan Peredaran 26,5 Kg Sabu	Narkoba
272	Kaleidoskop 2022: 10 Artis Terjerat Kasus Narkoba	Narkoba
273	Jelang Akhir Tahun Polda Metro Sita Ratusan Kilogram Narkoba, Ada yang Berupa Liquid Vape	Narkoba
274	Bareskrim Musnahkan 75 Kg Sabu dari 2 Oknum TNI Kurir Narkoba Jaringan Malaysia	Narkoba
275	Bongkar Peredaran Narkoba Jaringan Medan-Bali, BNNP Tangkap 8 Pelaku	Narkoba
276	TNI AL Tangkap Penyelundup 43 Paket Sabu di Lhokseumawe	Narkoba
277	Bareskrim Tangkap 4 Tersangka Kasus Peredaran Gelap Narkoba, 2 di Antaranya Oknum TNI	Narkoba
278	Bareskrim Tangkap 1 Buronan Kasus Narkoba di Kampung Ambon Jakarta Barat	Narkoba
279	Polisi Gagalkan Penyelundupan 40 Kg Sabu dalam Kotak Buah di Jaktim	Narkoba
280	Polres Badung Gulung 8 Pelaku Narkoba	Narkoba
281	Bareskrim Tangkap Akbar Antoni Penyelundup Sabu 179 Kg di Malaysia	Narkoba
282	Ganja Jadi 'Primadona' Kasus Narkotika di Sumedang	Narkoba
283	Bareskrim Minta Polda Sulsel Selidiki soal Viral Tersangka Narkoba Sebut Dapat "Backup" Polres	Narkoba
284	Bareskrim: 1,2 Ton Narkotika Disita Selama Periode Januari-Februari 2023	Narkoba
285	Polres Bandara Soekarno-Hatta Tangkap 19 Pengedar Narkoba sejak Februari, Salah Satunya WN Uganda	Narkoba
286	Bareskrim Musnahkan 373,23 Kg Sabu hingga 705 Butir Ekstasi dari 8 Kasus Peredaran Narkoba	Narkoba
287	Pengedar Narkoba di Pasangkayu Ditangkap, 16 Saset Sabu dari Sulteng Disita	Narkoba
288	Polda Metro Terima Penghargaan Muri atas Ungkap Kasus 277 Kg Sabu	Narkoba
289	Operasi Pekat Kota Batu 2023 Didominasi Penjualan Miras Ilegal	Narkoba
290	Operasi Pekat di Tulungagung, 80 Kilogram Obat Mercon-Ribuan Miras	Narkoba
291	Polisi Bongkar Gudang Penyimpanan Narkoba di Bekasi Senilai Rp 23 Miliar	Narkoba
292	Terjerat Narkoba, Artis Hud Filbert dkk Ditangkap di 3 Lokasi Berbeda	Narkoba
293	Rumah Mewah di Duren Sawit yang Digerebek Terkait Narkoba Ternyata Dikontrakkan	Narkoba
294	Gerebek Kampung Ambon, Polisi Tangkap Tiga Orang Diduga Pemakai Narkoba	Narkoba
295	Pelaku Narkoba Bebas Usai Bayar Rp 10 Juta, Propam Polda Sulsel Periksa Oknum Anggota Ditresnarkoba	Narkoba
296	Karyawan BUMN di Langkat Ditangkap gegara Jadi Pengedar Sabu	Narkoba
297	11 Napi Narkoba Rutan Negara Dites Urine, Gus Google Tunggu Jadwal	Narkoba
298	Napi Lapas Batam Kendalikan 12 Kg Sabu di Bandara Soetta Pakai HP	Narkoba
299	Pengedar 4,3 Kilogram Sabu di Konawe Sultra Ditangkap Polisi	Narkoba
300	Polisi Ungkap Modus Klasik Peredaran Narkoba di Cirebon	Narkoba