

# ANALISIS SENTIMEN OPINI PUBLIK TERKAIT JUDI ONLINE PADA PENGGUNA APLIKASI X MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MECHINE

Andrea Maulana<sup>1</sup>, Ade Yuliana<sup>2</sup>,

<sup>1,2</sup>Politeknik TEDC Bandung ; Jl. Politeknik Jl. Pesantren No.2, Cibabat, Kec. Cimahi Utara, Kota Cimahi, Jawa Barat 40513; telp (022) 6645951

Received: 12 Agustus 2024

Accepted: 5 Oktober 2024

Published: 12 Oktober 2024

## Keywords:

Naïve Bayes

Support Vector

Machine

Media sosial

Analisis Sentimen

**Abstrak.** Penelitian ini mengkaji sentimen masyarakat terkait perjudian online di Indonesia dengan menggunakan algoritma *Naive Bayes* dan *Support Vector Machine* (SVM). Data dikumpulkan dari platform X dan dianalisis melalui tahap-tahap prapemrosesan, klasifikasi, dan evaluasi. Hasil penelitian menunjukkan bahwa SVM memiliki tingkat akurasi sebesar 98%, lebih tinggi dibandingkan dengan *Naive Bayes* yang memiliki akurasi 93%. Ini menegaskan bahwa SVM lebih andal dalam mengklasifikasikan sentimen publik terhadap perjudian online. Temuan ini memberikan wawasan penting bagi pengambil kebijakan untuk memahami persepsi publik dan menyusun regulasi yang lebih tepat

## Correspondent Email:

[andreamaulana581@gmail.com](mailto:andreamaulana581@gmail.com)

**Abstract.** This research examines public sentiment regarding online gambling in Indonesia using the Naive Bayes algorithm and Support Vector Machine (SVM). Data is collected from the X platform and analyzed through preprocessing, classification, and evaluation stages. The research results show that SVM has an accuracy rate of 98%, higher than Naive Bayes which has an accuracy of 93%. This confirms that SVM is more reliable in classifying public sentiment towards online gambling. These findings provide important insights for policy makers to understand public perceptions and develop more appropriate regulations.

## 1. PENDAHULUAN

Salah satu isu sosial yang telah ada sejak lama adalah fenomena perjudian. Seiring dengan meningkatnya jumlah orang yang menggunakan alat komunikasi elektronik berbasis internet, perjudian berbasis teknologi, atau perjudian *online*, telah mengalami pertumbuhan dan perkembangan yang pesat [1]. Di Indonesia, aktivitas perjudian *online* semakin merambah ke seluruh Masyarakat Berbagai cara terus dilakukan untuk

memancing korban. Pelaku menggunakan kemajuan teknologi yang cepat untuk mengembangkan perbuatannya sekaligus menghindari terungkapnya sebagai akibat dari perjudian *online* [2]. Peneliti menganalisis sentimen data tweet pengguna X. Sebelumnya, terutama dengan menggunakan *algoritma Naive Bayes* dan algoritma SVM, seperti yang digunakan dalam penelitian yang dilakukan oleh Achmad Rizky Zulkarnain, Hafiz Irsyad [3] Mengenai klasifikasi opini masyarakat

mengenai pemblokiran judi *online*, terdapat penelitian dengan akurasi 61.84 dan skor F1 0.7590. Selain itu, terdapat penelitian Okta Nur Julianti, Nana Suarna, dan Willy Prihartono tentang Penerapan *Natural Language Processing* Pada Analisis Sentimen Perjudian *Online* Di Media Sosial Twitter [4]. Skor sentimen lebih negatif sebesar 75,8 dengan jumlah 722 data, sedangkan skor sentimen positif sebesar 7,7 dengan jumlah 73 data dan skor sentimen netral sebesar 16,5 dengan jumlah 157 data [5]. Berikutnya adalah penelitian yang dilakukan oleh Ditha Lozera Devi<sup>1</sup>, Penelitian AmaliaAnjani Arifiyanti dan Seftin Fitri Ana Wati dengan judul penelitian "Analisis Sentimen Ulasan Pengguna *Access Kai* Menggunakan Metode *Word2vec* Dan Algoritma *Svm*" memiliki nilai akurasi 81%. Penelitian ini diklasifikasikan menggunakan kernel sigmoid SVM. Algoritma menggunakan metode CBOV menghasilkan nilai akurasi 84%, dan klasifikasi menggunakan algoritma *sigmoid kernel SVM* menggunakan metode skip-gram menghasilkan nilai akurasi 84%. Tujuan dari penelitian ini adalah untuk melihat bagaimana masyarakat melihat perjudian *online* yang diungkapkan pada platform X dan untuk mengetahui bagaimana sentimen umum berkembang, apakah positif, negatif, atau netral.

## 2. TINJAUAN PUSTAKA

### 2.1 Sentimen

Sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, kalimat, atau pendapat. Polaritas mempunyai arti apakah teks yang ada dalam dokumen, kalimat, atau pendapat memiliki aspek positif atau negatif [6]. Saat ini, para peneliti banyak menggunakan analisis sentimen sebagai bidang penelitian dalam ilmu komputer. Analisis sentimen seringkali menggunakan jejaring sosial seperti Twitter untuk mengetahui persepsi Masyarakat [7].

### 2.2 Twitter (X)

*Twitter*, yang sekarang dikenal sebagai X, adalah platform media sosial di mana orang dapat berbagi pesan singkat, gambar, dan video. Platform ini telah berkembang menjadi Tempat di mana orang dari seluruh dunia berkumpul untuk membahas ide-ide baru dan

tren. Ulasan pada aplikasi X dibagi ke dalam suasana hati yang berbeda [8].

### 2.3 Data Mining

Data mining adalah kumpulan teknik yang digunakan secara otomatis untuk menggali dan memahami hubungan kompleks dalam kumpulan data yang sangat besar [9].

### 2.4 Algoritma *Naïve Bayes*

Metode klasifikasi *Naive Bayes* memanfaatkan prinsip probabilitas dan statistik yang dikembangkan oleh ilmuwan Inggris, *Thomas Bayes*. Pada intinya, metode ini menggunakan prinsip dasar *Teorema Bayes* untuk memprediksi kemungkinan kejadian di masa depan berdasarkan data dan pengalaman masa lalu [10].

### 2.5 Algoritma *Support Vector Machine*

Algoritma pembelajaran mesin yang dikenal sebagai *Support Vector Machine* (SVM) digunakan untuk menentukan kelas fitur data dan klasifikasi regresi [11]. Kemampuan SVM untuk menangani jumlah data yang sangat besar menjadikannya metode yang efektif dalam klasifikasi [12].

### 2.6 Klasifikasi

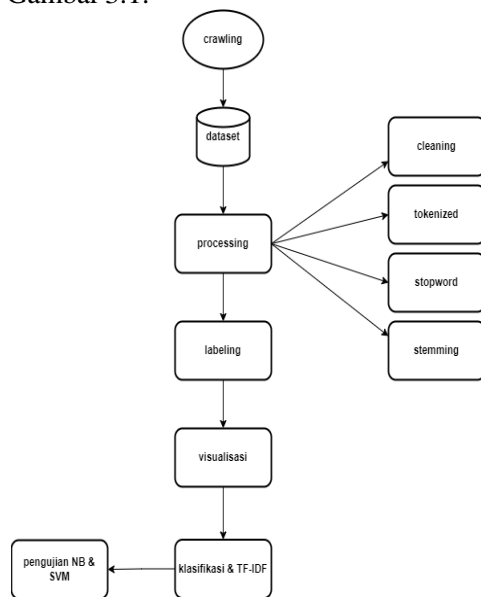
Metode klasifikasi digunakan untuk menemukan pola dalam kumpulan data yang mampu membedakan satu kelas data dari yang lain. dengan kelas data lainnya untuk menentukan objek mana yang termasuk dalam kategori tertentu berdasarkan perilaku dan atribut kelompok yang ditentukan. Teknik ini memungkinkan data baru disambiguasi dengan menggunakan hasil untuk memberikan seperangkat aturan [13].

### 2.7 TF-IDF

Metode TF-IDF menghitung berat setiap kata, dan ini adalah yang paling banyak digunakan. untuk pengambilan informasi. Cara ini diyakini efisien, sederhana, dan memberikan hasil yang akurat [14].

### 3. METODE PENELITIAN

Penulis menggunakan metodologi penelitian yang terdiri dari enam langkah: pencarian data, persiapan data (*pre-processing*), klasifikasi dengan metode *Naive Bayes* dan *SVM*, pembuatan matriks kebingungan (*confusion matrix*), serta evaluasi hasil. Gambar 3.1.



Gambar 3.1 Diagram Alur Penelitian

#### 3.1 Crawling Data

Data penelitian dikumpulkan dari 28 Desember 2023 hingga 29 Februari 2024 menggunakan *API* dan *Token X* untuk pengelolaan data. Hasil *crawling* data ditampilkan dalam Tabel 1.

Tabel 1 Data Set Hasil *Crawling*

No	Data Set
1	@RNHerdiansyah @PartaiSocmed buseet paraah seliweran mulu itu iklanjudi di TL kita
2	1. perokok 2. pemain judi online 3. pemain ML 4. suci anggun
3	@GOAL_ID report aja yuk nih @GOAL_ID atau blokir aja sekalian, kok anteng2 wae ditempelin iklan judi, omong kosong kalo mereka gak tahu, bangsat
4	@PartaiSocmed Pembatahan apapun ente g mainjudi,tapi ente ada ditempat judi!

#### 3.2 Pre-Processing

Proses ini mencakup berbagai langkah, seperti pembersihan data dari kesalahan atau ketidak konsistenan, transformasi data ke dalam format yang sesuai, pengisian nilai yang hilang, serta normalisasi atau standarisasi data untuk memastikan keseragaman. (Tabel 2-5).

Tabel 2 Hasil *Cleaning*

Sebelum	Sesudah
IKLAN JUDI BOLA SIALAN !! yang adaNikita Mirzani nya !	IKLAN JUDI BOLA SIALAN yang adaNikita Mirzani nya

Tabel 3 Hasil *Tokenizing*

Sebelum	Sesudah
@desmondwira Orang lebih senang judi dan menunjukkan kegagahantampung pisau jatuh Daripada berpikir warasDari 800 an saham di bursa, yg di lihat unvr	['@desmondwira', 'Orang', 'lebih', 'senang','judi', 'dan', 'menunjukkan', 'kegagahan', 'tampung', 'pisau', 'jatuh', 'Daripada', 'berpikir', 'waras', 'Dari', '800', 'an', 'saham', 'di', 'bursa,', 'yg', 'di', 'lihat','unvr']

Tabel 3 Hasil *Stopword*

Sebelum	Sesudah
@praburevolusi @donnygunawan8 2 Padahal smua BA Judionline ngumpulnya disitu tuh Hahahahaha	@praburevolusi @donnygunawan8 2Padahal BA Judi online ngumpulnya Hahahahaha

Tabel 5 Hasil *Stemming*

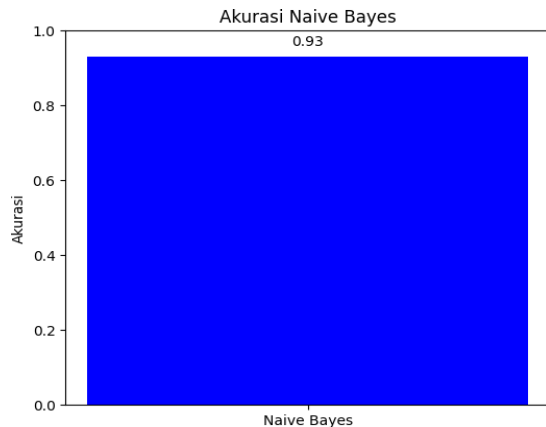
Sebelum	Sesudah
@fitridefi Kalo bapaknya nonton youtube tentang judislot gimana?	@fitridefi Kalau bapaktonton youtube tentang judi slot bagaimana?

### 3.3 Labelling

Labelling adalah proses pemberian label atau kategori pada suatu teks berdasarkan perasaan atau makna yang terkandung di dalamnya. dari memberi label atau kategori pada suatu teks berdasarkan perasaan atau makna yang terkandung di dalamnya. Umumnya, teks dari artikel, komentar, atau media sosial media diberi label positif ,negatif, atau netral diberi label positif ,negatif, atau netral.

### 3.4 Klasifikasi *Naïve Bayes*

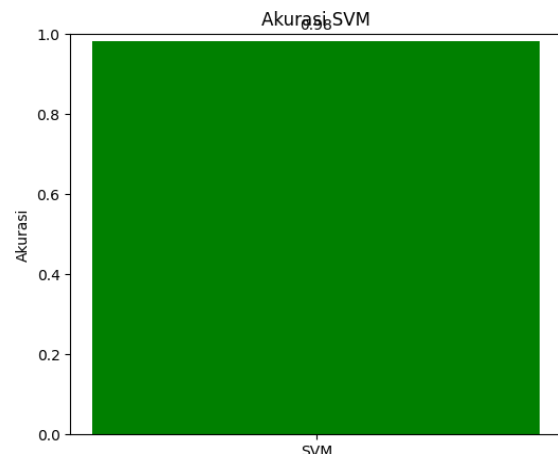
Gambar tersebut menunjukkan visualisasi akurasi model *Naïve Bayes*. Grafik ini memiliki judul "Akurasi *Naïve Bayes*" dan menampilkan sebuah bilah alat berwarna biru yang menggambarkan tingkat akurasi model *Naïve Bayes* sebesar 93% (atau 0.93 pada skala 1). Sumbu vertikal dilabeli "Akurasi" dengan rentang nilai dari 0 hingga 1.



Gambar 3.2 Hasil Klasifikasi *Naïve Bayes*

### 3.5 Klasifikasi *Support Vector Machine*

Gambar ini merupakan visualisasi akurasi model *Support Vector Machine* (SVM). Grafik tersebut memiliki judul "Akurasi SVM" dan menampilkan sebuah bilah pisau berwarna hijau yang menunjukkan akurasi model SVM. Nilai akurasi yang ditampilkan pada bilah alat dengan angka 0.98 atau 98% pada Skala 1. Sumbu vertikal dilabeli "Akurasi" dengan nilai berkisar dari 0 hingga 1, yang menunjukkan performa model yang sangat tinggi dalam klasifikasi.



Gambar 3.3 Hasil Klasifikasi SVM

### 3.6 *Confusion Matrix*

Matriks konfusi adalah tabel untuk mengevaluasi kinerja model klasifikasi. Matriks ini menunjukkan bagaimana model membuat prediksi dan seberapa sering hasilnya benar atau salah. Matriks kebingungan membantu anda menghitung berbagai metrik seperti perolehan, skor F1, presisi, dan presisi, sehingga memberi Anda gambaran yang lebih detail tentang performa model Anda dibandingkan dengan menggunakan presisi saja.

### 3.7 Evaluasi

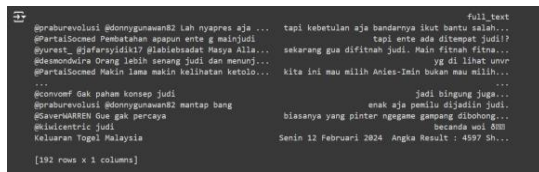
Evaluasi dalam pembelajaran mesin adalah proses mengevaluasi kinerja suatu model. Aspek penting dari evaluasi model mencakup akurasi, yang mengukur proporsi ketepatan menunjukkan persentase prediksi positif yang benar dari jumlah total prediksi positif. ingatan atau kepekaan. Menunjukkan persentase kasus positif yang terdeteksi dengan benar dari seluruh kasus positif. Skor F1 adalah rata-rata harmonik antara presisi dan perolehan.

## 4. HASIL DAN PEMBAHASAN

### 4.1 *Crawling*

Data yang di terdiri dari 630 sampel yang diambil dalam bahasa Indonesia selama jangka waktu tertentu. Data ini digunakan untuk berbagai kategori penelitian.

Gambar berikut menunjukkan contoh pengumpulan data.



Gambar 4.1 hasil *Crawling*

## 4.2 Pre-Processing

Dalam data mining, Proses persiapan data yang meliputi *preprocessing* mencakup pembersihan, transformasi, dan penyaringan. Tujuannya adalah untuk memastikan bahwa data tersebut siap dan relevan sebelum melanjutkan ke tahap pemrosesan lebih lanjut. *Cleaning* Proses ini mencakup langkah-langkah seperti menghapus duplikat, mengisi nilai yang hilang, memperbaiki kesalahan format data, dan menghapus outlier dan data yang tidak relevan. Pembersihan data membantu meningkatkan akurasi dan keandalan hasil analisis dan model yang Anda buat.

Full Text  
@RHHerdiansyah @PartaSoMed buset parah seliweran mulu itu iklan judi di TL kita  
1. perokok 2. pemain judi online 3. pemain ML 4. suci anggun  
@GOAL ID report aja yuk nih @GOAL ID atau blokir aja sekalian, kok anteng2 wae ditempelin iklan judi, omong kosong kalo mereka gak tahu, bangsat  
@PartaSoMed Pembatahan apapun ente g mainjudi tapi ente ada ditempat judi!

Gambar 4.2 sebelum proses *Cleaning*

Full Text  
buset parah seliweran mulu itu iklan judi di TL kita  
perokok pemain judi online pemain ML suci anggun  
report aja yuk atau blokir aja sekalian kok anteng2 wae ditempelin iklan judi omong kosong kalo mereka gak tahu bangsat  
Pembatahan apapun ente g mainjudi tapi ente ada ditempat judi!

Gambar 4.3 setelah proses *Cleaning*

### 4.2.1 Tokenizing

Proses membagi teks menjadi unit-unit yang lebih kecil disebut "tokenisasi." Bergantung pada konteks dan tujuan analisis, token dapat berupa kata, huruf, atau bahkan subkata analisis.

Full Text  
buset parah seliweran mulu itu iklan judi di TL kita  
perokok pemain judi online pemain ML suci anggun  
report aja yuk atau blokir aja sekalian kok anteng2 wae ditempelin iklan judi omong kosong kalo mereka gak tahu bangsat  
Pembatahan apapun ente g mainjudi tapi ente ada ditempat judi!

Gambar 4.4 sebelum di *Tokenizing*

Full Text  
[buset', 'parah', 'seliweran', 'mulu', 'itu', 'iklan', 'judi', 'di', 'TL', 'kita']  
[perokok', 'pemain', 'judi', 'online', 'pemain', 'ML', 'suci', 'anggun']  
[report', 'aja', 'yuk', 'atau', 'blokir', 'aja', 'sekalian', 'kok', 'anteng2', 'wae', 'ditempelin', 'iklan', 'judi', 'omong', 'kosong', 'kalo', 'mereka', 'gak', 'tahu', 'bangsat']  
[Pembatahan', 'apapun', 'ente', 'g', 'mainjudi', 'tapi', 'ente', 'ada', 'ditempat', 'judi']

Gambar 4.5 setelah *Tokenizing*

### 4.2.2 Stopword

Kata-kata umum dalam suatu bahasa, seperti "dan", "atau", "desu", dan "kore", yang sering muncul namun biasanya tidak mempunyai arti penting dalam analisis teks.

Full Text  
[buset', 'parah', 'seliweran', 'mulu', 'itu', 'iklan', 'judi', 'di', 'TL', 'kita']  
[perokok', 'pemain', 'judi', 'online', 'pemain', 'ML', 'suci', 'anggun']  
[report', 'aja', 'yuk', 'atau', 'blokir', 'aja', 'sekalian', 'kok', 'anteng2', 'wae', 'ditempelin', 'iklan', 'judi', 'omong', 'kosong', 'kalo', 'mereka', 'gak', 'tahu', 'bangsat']  
[Pembatahan', 'apapun', 'ente', 'g', 'mainjudi', 'tapi', 'ente', 'ada', 'ditempat', 'judi']

Gambar 4.6 sebelum di *Stopword*

Full Text

[buset', 'parah', 'iklan', 'judi', 'TL', 'kita']  
[perokok', 'pemain', 'judi', 'online', 'ML', 'suci', 'anggun']  
[report', 'blokir', 'iklan', 'judi', 'omong', 'kosong', 'bangsat']  
[Pembatahan', 'mainjudi', 'ditempat', 'judi']

Gambar 4.7 setelah di *Stopword*

### 4.2.3 Stemming

Proses mereduksi sebuah kata menjadi bentuk dasarnya atau "akar" dengan menghilangkan akhiran yang sama. Misalnya, kata "running", "runner", dan "run" semuanya dapat diubah menjadi "run" sebagai bentuk dasarnya.

Full Text  
[buset', 'parah', 'iklan', 'judi', 'TL', 'kita']  
[perokok', 'pemain', 'judi', 'online', 'ML', 'suci', 'anggun']  
[report', 'blokir', 'iklan', 'judi', 'omong', 'kosong', 'bangsat']  
[Pembatahan', 'mainjudi', 'ditempat', 'judi']

Gambar 4.8 sebelum di *Stemming*

Full Text  
[buset', 'parah', 'iklan', 'judi', 'TL', 'kita']  
[rokok', 'main', 'judi', 'online', 'ML', 'suci', 'anggun']  
[report', 'blokir', 'iklan', 'judi', 'omong', 'kosong', 'bangsat']  
[batuahan', 'mainjudi', 'tempat', 'judi']

Gambar 4.9 setelah di *Stemming*

### 4.2.4 Labelling

Labelling yaitu proses analisis teks untuk mengklasifikasikan konten judi *online*, mulai dari pembersihan teks, pemecahan menjadi token, hingga penentuan label "Judi" atau "Nonjudi".

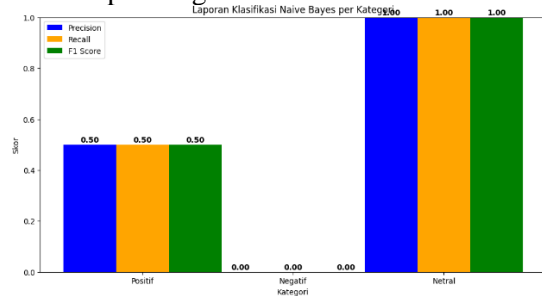
Full Text	Cleaned Text	Tokenized Text	Text Without Stopwords	Stemmed Text	Sentiment
@praburevolusi @donnygunawan82 Lah nyapra aja judi tapi kabetulan aja bandarnya ikut bantu salah...	lah nyapra aja judi tapi kabetulan aja bandarnya ikut bantu salah...	lah nyapra aja judi tapi kabetulan aja bandarnya ikut bantu salah...	lah nyapra aja judi tapi kabetulan aja bandarnya ikut bantu salah...	lah nyapra aja judi tapi kabetulan aja bandarnya ikut bantu salah...	Positif
@PartaSoMed Pembatahan apapun ente g mainjudi tapi ente ada ditempat judi!	apapun ente g mainjudi tapi ente ada ditempat judi!	apapun ente g mainjudi tapi ente ada ditempat judi!	apapun ente g mainjudi tapi ente ada ditempat judi!	apapun ente g mainjudi tapi ente ada ditempat judi!	Netral
@Gyrest @djarfaryid137 @labiehadat Masya Alla... seharang gua difitnah judi. Main fitnah fitna...	seharang gua difitnah judi. Main fitnah fitna...	seharang gua difitnah judi. Main fitnah fitna...	seharang gua difitnah judi. Main fitnah fitna...	seharang gua difitnah judi. Main fitnah fitna...	Positif
@donnygunawan82 Orang jodoh senang judi dan menang... kita ini mau milih Anies-Inin bukan mau milih...	orang jodoh senang judi dan menang... kita ini mau milih Anies-Inin bukan mau milih...	orang jodoh senang judi dan menang... kita ini mau milih Anies-Inin bukan mau milih...	orang jodoh senang judi dan menang... kita ini mau milih Anies-Inin bukan mau milih...	orang jodoh senang judi dan menang... kita ini mau milih Anies-Inin bukan mau milih...	Netral
@donnygunawan82 @SaverWARREN Gua gak percaya @Luiscentric judi kaluaran Topol Malaysia Senin 12 Februari 2024 Angka Result : 4897 Sh...	gua gak percaya @Luiscentric judi kaluaran Topol Malaysia Senin 12 Februari 2024 Angka Result : 4897 Sh...	gua gak percaya @Luiscentric judi kaluaran Topol Malaysia Senin 12 Februari 2024 Angka Result : 4897 Sh...	gua gak percaya @Luiscentric judi kaluaran Topol Malaysia Senin 12 Februari 2024 Angka Result : 4897 Sh...	gua gak percaya @Luiscentric judi kaluaran Topol Malaysia Senin 12 Februari 2024 Angka Result : 4897 Sh...	Positif
@donnygunawan82 @SaverWARREN Gua gak percaya @Luiscentric judi kaluaran Topol Malaysia Senin 12 Februari 2024 Angka Result : 4897 Sh...	gua gak percaya @Luiscentric judi kaluaran Topol Malaysia Senin 12 Februari 2024 Angka Result : 4897 Sh...	gua gak percaya @Luiscentric judi kaluaran Topol Malaysia Senin 12 Februari 2024 Angka Result : 4897 Sh...	gua gak percaya @Luiscentric judi kaluaran Topol Malaysia Senin 12 Februari 2024 Angka Result : 4897 Sh...	gua gak percaya @Luiscentric judi kaluaran Topol Malaysia Senin 12 Februari 2024 Angka Result : 4897 Sh...	Netral

Gambar 4.10 Hasil Labelling

### 4.3 Evaluasi

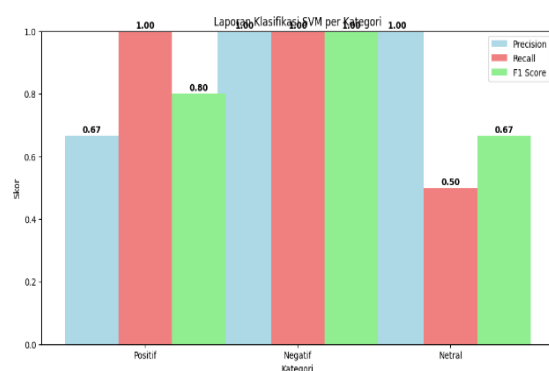
#### 4.3.1 Hasil Performance

Pengambilan data pada platform X untuk analisis perjudian *online* dengan algoritma *Naive Bayes* dan dukungan *mechine vector* menggunakan 630 kumpulan data. Hasil dari *Google Collab* menunjukkan bahwa algoritma *Naive Bayes* dan SVM dapat meningkatkan analisis persaingan.



Gambar 4.11 Hasil Performance NB

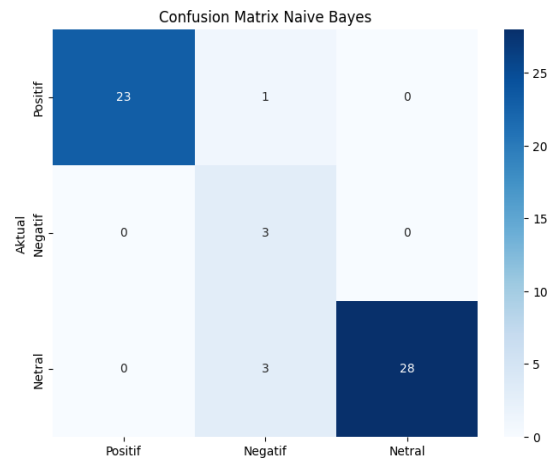
Recall, dan F1 Score yang sama, yaitu 0.50. Kategori "Negatif" menunjukkan kinerja yang sangat rendah dengan nilai nol pada semua metrik, menandakan bahwa model tidak mampu mengidentifikasi kategori ini dengan baik. Sebaliknya, pada kategori "Netral," model mencapai kinerja sempurna dengan nilai 1.00 pada semua metrik, menunjukkan kemampuan yang sangat baik dalam mengklasifikasikan data ke dalam kategori ini. Secara keseluruhan, model *Naive Bayes* lebih unggul dalam mendeteksi kategori "Netral," tetapi lemah dalam mengenali kategori "Negatif."



Gambar 4.12 Hasil Performance SVM

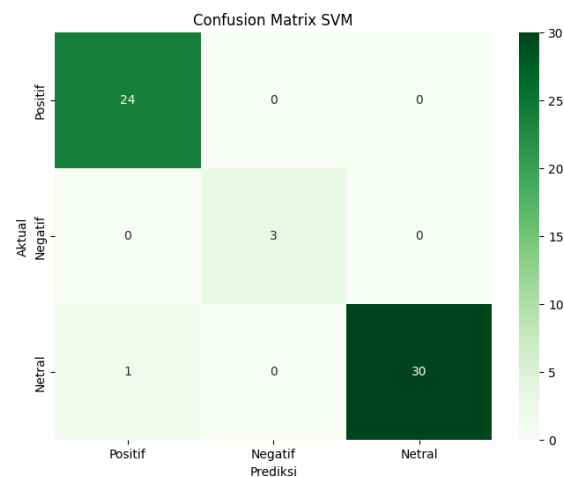
Gambar tersebut menunjukkan laporan klasifikasi model SVM untuk tiga kategori: Positif, Negatif, dan Netral. Pada kategori Positif dan Negatif, *Precision* mencapai 1.00, dengan F1 Score masing-masing 1.00 dan 0.80. Namun, performa pada kategori Netral

lebih rendah, dengan *Precision* dan F1 Score 0.67, serta Recall 0.50. Secara keseluruhan, SVM tampil baik pada kategori Positif dan Negatif, tetapi kurang optimal pada kategori Netral.



Gambar 4.13 Confusion Matrix Naive Bayes

Tampilan ini adalah matriks konfusi untuk model *Naive Bayes*. Model tersebut memprediksi dengan tepat 23 kasus positif, 3 kasus negatif, dan 28 kasus netral. Terdapat kesalahan prediksi sebesar 0 kasus positif dianggap netral dan 3 kasus netral dianggap negatif. Model ini berkinerja sangat baik pada data netral, namun memiliki beberapa kesalahan pada data positif dan negatif.



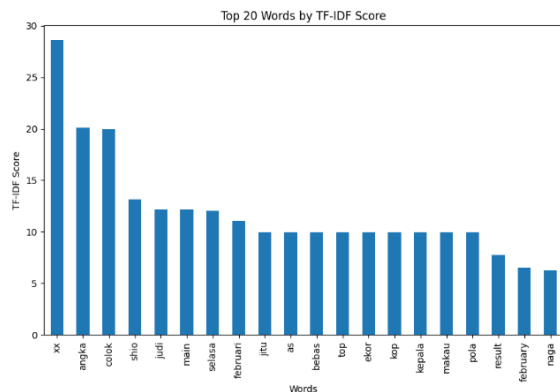
Gambar 4.14 Confusion Matrix SVM

Tampilan ini adalah matriks konfusi untuk model *support vector machine* (SVM). Model tersebut memprediksi dengan tepat 24 kasus positif, 3 kasus negatif, dan 30 kasus netral.



Terjadi kesalahan prediksi dengan 0 kasus positif diprediksi netral dan 0 kasus netral diprediksi negatif. Model ini berkinerja baik pada data positif dan netral, namun membuat beberapa kesalahan pada kategori netral dan negatif.

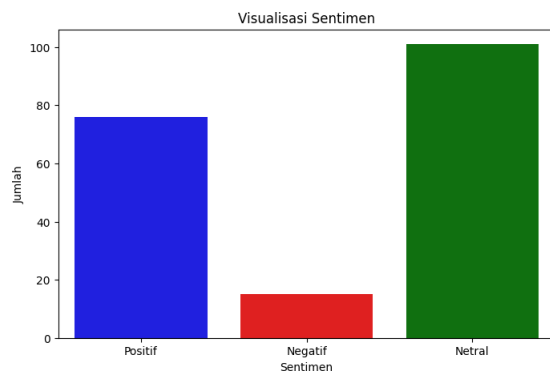
### 4.3.2 TF-IDF



Gambar 4.15 Visualisasi TF-IDF

Gambar tersebut menampilkan grafik batang yang menunjukkan 20 kata teratas berdasarkan skor TF-IDF (*Term Frequency - Inverse Document Frequency*). Skor TF-IDF mengukur pentingnya kata dalam sebuah dokumen relatif terhadap kumpulan dokumen. Kata dengan skor tertinggi adalah "xx," diikuti oleh kata-kata seperti "angka," "colok," dan "shio." Kata-kata ini dianggap paling signifikan atau relevan dalam konteks dokumen yang dianalisis. Semakin tinggi skor TF-IDF, semakin besar kontribusi kata tersebut dalam membedakan dokumen dalam kumpulan data.

### 4.3.3 Visualisasi Sentimen

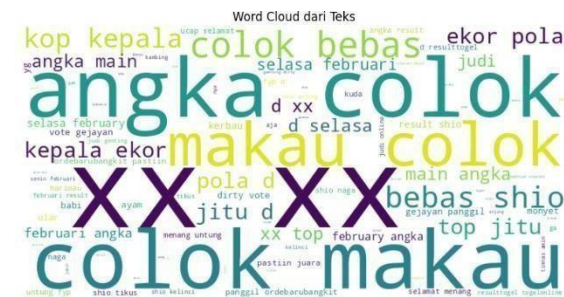


Gambar 4.16 Gambar Diagram Visualisasi sentimen

Grafik di atas menunjukkan sebaran opini masyarakat terhadap suatu isu. Mayoritas opini bersifat netral sekitar 100, disusul sentimen positif sekitar 80, dan sentimen paling banyak negatif sekitar 20. Hal ini menunjukkan bahwa opini masyarakat lebih netral dan lebih cenderung positif di bandingkan negatif.

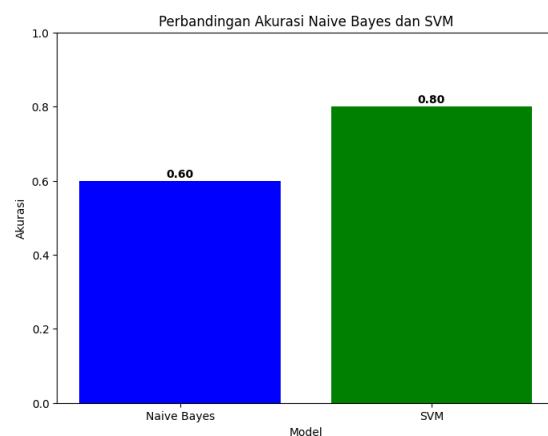
#### 4.3.4 Word Cloud

Gambar ini adalah awan kata yang menunjukkan kata-kata yang paling sering digunakan dalam teks besar seperti “colok”, “xx”, “x”, dan “angka” lebih sering atau lebih jelas muncul dalam teks. *Word Cloud* membantu Anda memvisualisasikan frekuensi kata dan mengidentifikasi topik penting.



Gambar 4.17 Gambar *Word Cloud*

#### 4.4.4 Hasil Perbandingan Akurasi NB dan SVM



Gambar 4.18 Gambar Diagram Hasil Akurasi

Gambar ini membandingkan akurasi dua model pembelajaran mesin: *Naive Bayes* (biru) dengan akurasi 60% dan SVM (hijau) dengan akurasi 80%. Grafik menunjukkan bahwa SVM memiliki akurasi lebih tinggi dibandingkan *Naive Bayes*.

## 5. KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

1. Penelitian ini berhasil menunjukkan bahwa algoritma *Support Vector Machine* (SVM) lebih unggul dalam hal akurasi dibandingkan dengan *Naive Bayes*.
2. Data yang dianalisis menunjukkan bahwa sebagian besar opini masyarakat terkait perjudian online bersifat netral.
3. Hasil ini memberikan wawasan berharga bagi para pembuat kebijakan dalam mengambil langkah yang lebih tepat dalam menghadapi isu perjudian *online*.

### 5.2 Saran

1. Meskipun SVM unggul, disarankan untuk mengeksplorasi algoritma lain seperti *Random Forest* atau *Deep Learning* untuk meningkatkan akurasi klasifikasi sentimen.
2. Penelitian lanjutan bisa difokuskan pada analisis sentimen berdasarkan kelompok demografi tertentu, seperti usia atau wilayah, untuk memperoleh wawasan yang lebih spesifik.
3. Peningkatan jumlah data dan penggunaan model *hybrid* juga bisa dipertimbangkan guna mengurangi bias dan meningkatkan generalisasi hasil.

## DAFTAR PUSTAKA

- [1] R. Sepatia, T. R. Zarzani, and M. Purba, "Analisis yuridis pertanggungjawaban pidana bagi pembuat website yang dipergunakan untuk perjudian online (Analisis Putusan No. 852/Pid. Sus/2020/PN. Mdn)," *Jurnal Rectum*, vol. 4, no. 2, pp. 430-442, Jul. 2022.
- [2] PPATK, "Aliran dana terkait judi terus dipantau," *Pusat Pelaporan dan Analisis Transaksi Keuangan*,
- [3] R. Antonius, A. R. Zulkarnain, and H. Irsyad, "Pendekatan TF-IDF, SMOTE, dan SVM dalam Klasifikasi Sentimen Masyarakat terhadap Pemblokiran Judi Online," *Buletin Ilmiah Informatika Teknologi*, vol. 2, no. 3, pp. 115-122, 2024.
- [4] O. N. Julianti, N. Suarna, and W. Prihartono, "Penerapan Natural Language Processing pada Analisis Sentimen Judi Online di Media Sosial Twitter," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 3, pp. 2936-2941, Jun. 2024.
- [5] D. L. Devi, A. A. Arifiyanti, dan S. F. A. Wati, "Analisis Sentimen Ulasan Pengguna Access By Kai Menggunakan Metode Word2vec Dan Algoritma Svm," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Agu 2024, doi: 10.23960/jitet.v12i3.4892.
- [6] Nugroho, Dirjen, Chrisnanto, YH, & Wahana, A. (2016). Analisis Sentimen pada Jasa Ojek Online Menggunakan Metode *Naïve Bayes*. *Prosiding SNST ke-7 Fakultas Teknik Universitas Wahid Hasyim Semarang*, 156-161. ISBN 978-602-99334-5-1.
- [7] S. Juanita, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID- 19 Menggunakan Metode Naïve Bayes," *Jurnal Media Informatika Budidarma*, vol. 5, no. 1, pp. 157-163, Jan. 2021, doi:10.30865/mib.v5i1.2604.
- [8] S. Styawati, N. Hendrastuty, A. R. Isnain, and A. Y. Rahmadhani, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *Jurnal Informatika: Jurnal pengembangan IT (JPIT)*, vol. 6, no. 3, pp. 150-155, Sep. 2021
- [9] I. Ernawati, "Data Mining Sebagai Salah Satu Solusi Strategi Bisnis," *INFORMATIK Jurnal Ilmu Komputer*, vol. 14, no. 1, pp. 9-15, Apr. 2018.
- [10] A. F. Watratan, A. P. B., and D. Moeis, "Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia," *Journal of Applied Computer Science and Technology (JACOST)*, vol. 1, no. 1, pp. 7-14, 2020.
- [11] RAntonius, A. R. Zulkarnain, and H. Irsyad, "Pendekatan TF-IDF, SMOTE, dan SVM dalam Klasifikasi Sentimen Masyarakat terhadap Pemblokiran Judi Online," *Buletin Ilmiah Informatika Teknologi*, vol. 2, no. 3, pp. 115-122, Mei 2024, doi: 10.58369/biit.v2i3.65.
- [12] A. S. Styawati, N. Hendrastuty, A. R. Isnain, and Y. Rahmadhani, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *Jurnal Informatika: Jurnal Pengembangan IT (JPIT)*, vol. 6, no. 3, pp. 1-10, Sep. 2021, ISSN: 2477-5126, e-ISSN: 2548- 9356.
- [13] E. Nurlia and U. Enri, "Penerapan Fitur Seleksi Forward Selection untuk Menentukan Kematian Akibat Gagal Jantung Menggunakan Algoritma C4.5," *Jurnal Teknik Informatika Musirawas (JUTIM)*, vol. 6, no. 1, pp. 42-50, Jun. 2021.



- [14]A. Azis Maarif, "Penerapan Algoritma TF-IDF untuk Pencarian Karya Ilmiah," presented at the Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Semarang, 2021. [Online]. Available:  
<https://core.ac.uk/display/35380748>