

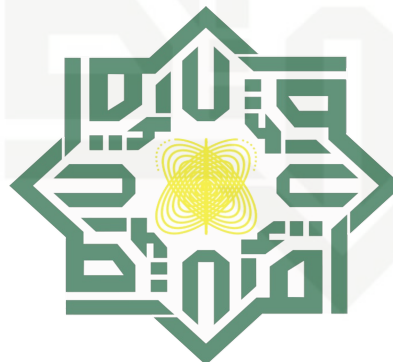
# ANALISIS SENTIMEN LIGA INDONESIA MENGGUNAKAN ALGORITMA *MACHINE LEARNING* PADA PLATFORM X DAN YOUTUBE

## TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat  
untuk Memperoleh Gelar Sarjana Komputer pada  
Program Studi Sistem Informasi

Oleh:

**MAHYUDA IRWANDA**  
**12150315127**



UIN SUSKA RIAU

State Islamic University of Sultan Syarif Kasim Riau

FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU  
PEKANBARU

2025

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.  
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LEMBAR PERSETUJUAN

### ANALISIS SENTIMEN LIGA INDONESIA MENGGUNAKAN ALGORITMA MACHINE LEARNING PADA PLATFORM X DAN YOUTUBE

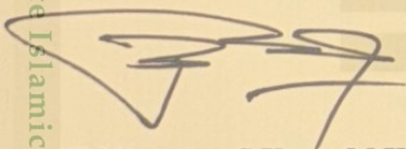
#### TUGAS AKHIR

Oleh:

**MAHYUDA IRWANDA**  
**12150315127**

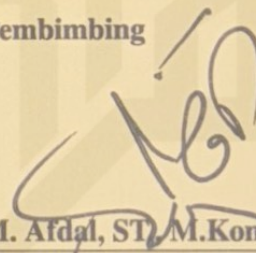
Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir  
di Pekanbaru, pada tanggal 03 Juli 2025

Ketua Program Studi



**Eki Saputra, S.Kom., M.Kom.**  
**NIP. 198307162011011008**

Pembimbing



**M. Afdal, ST, M.Kom.**  
**NIP. 198803282023211017**

UIN SUSKA RIAU



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PENGESAHAN

ANALISIS SENTIMEN LIGA INDONESIA MENGGUNAKAN  
ALGORITMA MACHINE LEARNING PADA PLATFORM X  
DAN YOUTUBE

TUGAS AKHIR

Oleh:


**MAHYUDA IRWANDA**  
**12150315127**

Telah dipertahankan di depan sidang dewan penguji  
sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer  
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau  
di Pekanbaru, pada tanggal 24 Juni 2025

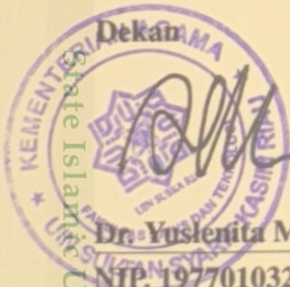
Pekanbaru, 03 Juli 2025

Mengesahkan,

Ketua Program Studi



**Eki Saputra, S.Kom., M.Kom.**  
**NIP. 1983071620111011008**



**Dr. Yuslenita Muda, S.Si., M.Sc.**  
**NIP. 197701032007102001**

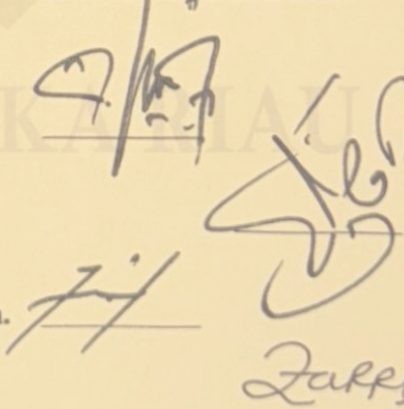
DEWAN PENGUJI:

Ketua : Arif Marsal, Lc., MA.

Sekretaris : M. Afdal, ST., M.Kom.

Anggota 1 : Dr. Rice Novita, S.Kom., M.Kom.

Anggota 2 : Zarnelly, S.Kom., M.Sc.



## LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum, dengan ketentuan bahwa hak cipta ada pada peneliti. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan atas izin peneliti dan harus dilakukan mengikuti kaedah dan kebiasaan ilmiah serta menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin tertulis dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan dapat meminjamkan Tugas Akhir ini untuk anggotanya dengan mengisi nama, tanda peminjaman dan tanggal pinjam pada *form* peminjaman.

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan di dalam daftar pustaka.

Pekanbaru, 03 Juli 2025

Yang membuat pernyataan,

**MAHYUDA IRWANDA**  
**NIM. 12150315127**

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



## LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

*Dengan menyebut nama Allah yang maha pengasih lagi maha penyayang*

*Assalamu'alaikum Warahmatullahi Wabarakatuh.*

*Alhamdulillah Rabbil 'Alamiin, segala puji bagi Allah Subhanahu Wa Ta'ala sebagai bentuk rasa syukur atas segala nikmat yang telah diberikan tanpa ada kekurangan sedikitpun. Shalawat beserta salam tak lupa pula kita ucapkan kepada Nabi Muhammad Shallallahu 'Alaihi Wa Sallam dengan mengucapkan Allahumma Sholli'ala Sayyidina Muhammad Wa'ala Ali Sayyidina Muhammad. Semoga kita semua selalu senantiasa mendapat syafaat-Nya di dunia maupun di akhirat, Aamiin Ya Rabbal'aalamiin.*

Tugas Akhir ini peneliti dedikasikan kepada orang tua tercinta Ayanhanda Mulyadi dan Ibunda Suyati yang telah berjuang tanpa lelah demi mengantarkan peneliti mencapai pendidikan tinggi. Berkat doa dan kasih sayangmu, anakmu telah berhasil memperoleh gelar sarjana seperti yang engkau harapkan. Tiada apapun di dunia ini yang dapat membalas semua jasa-jasa dan pengorbananmu. Peneliti juga berterima kasih yang tak terhingga kepada Mahdiah Fitri selaku Kakak yang selalu memberikan dukungan, semangat, pelajaran, serta pemahaman mengenai indahny kehidupan yang damai sebagai saudara.

Kepada Bapak dan Ibu Dosen Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau yang telah memberikan ilmu pengetahuan bermanfaat, pengalaman berharga, dan kebaikan yang tulus selama perkuliahan, peneliti ucapkan terima kasih banyak dan semoga menjadi amal jariyah. *Aamiin Ya Rabbal'aalamiin.*

Sahabat-sahabat terdekat yang tidak bisa peneliti sebutkan satu-persatu dan pastinya juga teman-teman seperjuangan, terima kasih berkat kalian masa perkuliahan menjadi lebih bermakna dan menyenangkan semoga di masa mendatang kita bisa bertemu lagi dalam keadaan yang lebih baik.

*Wassalamu'alaikum Warahmatullahi Wabarakatuh.*

#### Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## KATA PENGANTAR

*Assalamu'alaikum Warahmatullahi Wabarakatuh.*

*Alhamdulillah Rabbil 'Alamin*, puji syukur peneliti ucapkan kepada Allah *Subhanahu Wa Ta'ala* yang telah memberikan rahmat dan hidayah-Nya sehingga peneliti bisa menyelesaikan Tugas Akhir ini dengan hasil yang baik. *Shalawat* serta salam juga senantiasa dihadiahkan kepada Nabi Muhammad *Shallallahu 'Alaihi Wa Sallam* dengan mengucapkan *Allahumma Shalli'ala Sayyidina Muhammad Wa'ala Ali Sayyidina Muhammad*.

Laporan ini disusun sebagai syarat untuk memperoleh gelar sarjana dan sebagai pembelajaran akademis maupun spiritual. Peneliti mengucapkan terima kasih kepada semua pihak yang membantu dalam segala proses penelitian yang telah peneliti lakukan baik berupa materi, motivasi dan doa. Untuk itu, pada kesempatan ini peneliti mengucapkan terima kasih yang sangat mendalam kepada:

1. Ibu Prof. Dr. Hj. Leny Nofianti MS., SE., M.Si., Ak., CA selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Ibu Dr. Yuslenita Muda, S.Si., M.Sc selaku Dekan Fakultas Sains dan Teknologi.
3. Bapak Eki Saputra, S.Kom., M.Kom selaku Ketua Program Studi Sistem Informasi.
4. Ibu Siti Monalisa, ST., M.Kom selaku Sekretaris Program Studi Sistem Informasi.
5. Bapak Tengku Khairil Ahsyar, S.Kom., M.Kom selaku Kepala Laboratorium Program Studi Sistem Informasi.
6. Ibu Zarnelly, S.Kom., M.Sc selaku Dosen Pembimbing Akademik sekaligus Penguji II yang selalu memberikan arahan selama masa perkuliahan serta ilmu pengalaman dan waktu yang dimiliki untuk membantu peneliti menyelesaikan Tugas Akhir ini.
7. Bapak M. Afdal, ST., M.Kom sebagai dosen Pembimbing Tugas Akhir yang telah meluangkan waktu, tenaga, dan pikiran dalam membimbing peneliti hingga peneliti dapat menyelesaikan Laporan Tugas Akhir ini.
8. Bapak Arif Marsal, Lc., MA selaku Ketua Sidang yang telah meluangkan waktunya untuk melaksanakan sidang Tugas Akhir peneliti dan memberikan arahan tambahan terkait penelitian ini.
9. Ibu Dr. Rice Novita, S.Kom., M.Kom selaku Dosen Penguji I yang telah meluangkan waktunya dan membantu memberikan arahan tambahan terkait penelitian Tugas Akhir.



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

10. Bapak dan Ibu Dosen Program Studi Sistem Informasi yang telah banyak memberikan ilmunya kepada peneliti selama masa perkuliahan.
11. Teristimewa Ayahanda Muliyadi dan Ibunda Suyati, kedua sosok yang paling peneliti sayangi. Kasih sayang, dukungan, dan doa tanpa henti dari ayahanda dan ibunda adalah kekuatan terbesar peneliti dalam menyelesaikan Tugas Akhir ini. Setiap kata semangat dari ayahanda dan setiap doa dari ibunda selalu menjadi penyemangat peneliti untuk terus berjuang.
12. Kakak peneliti yaitu Mahdiah Fitri yang memberikan motivasi hingga peneliti dapat menyelesaikan Tugas Akhir ini.
13. Saudari Mila yang telah banyak memberikan motivasi, dukungan, dan semangat kepada peneliti.
14. Teman-teman seperjuangan SIF D 21 yang telah berjuang bersama menuntut ilmu dalam mengejar mimpi.
15. Semua pihak yang ikut serta dalam proses penelitian yang tidak dapat peneliti sebutkan satu-persatu.

Terima kasih yang sangat mendalam, semoga segala kebaikan yang telah diberikan menjadi ladang pahala serta mendapatkan balasan dari Allah *Subhanahu Wa Ta'ala*. Semoga kita semua selalu mendapatkan kebahagiaan dan kesehatan, *Aamin Ya Rabbal'aalamiin*. Peneliti menyadari bahwa penulisan Tugas Akhir ini masih banyak terdapat kekurangan dan jauh dari kata sempurna. Untuk itu kritik dan saran atau pertanyaan dapat diajukan melalui *e-mail* 12150315127@students.uin-suska.ac.id. Semoga laporan ini bermanfaat bagi kita semua. Akhir kata peneliti ucapkan terima kasih.

*Wassalamu'alaikum Warahmatullahi Wabarakaatuh.*

Pekanbaru, 03 Juli 2025

Peneliti,

**MAHYUDA IRWANDA**  
**NIM. 12150315127**



Medan, 20 May 2025

No : 705/BITS/LOA/V/2025

Lamp : -

Hal : Penerimaan Naskah Publikasi Ilmiah

Kepada Yth,

Bapak/Ibu **Mahyuda Irwanda**

Di Tempat

Terimakasih telah mengirimkan artikel ilmiah untuk diterbitkan pada **BUILDING OF INFORMATICS, TECHNOLOGY AND SCIENCE (BITS)** ISSN 2684-8910 (Print), ISSN 2685-3310 (Online), dengan judul:

### **Analisis Sentimen Masyarakat Terhadap Liga Indonesia Menggunakan Algoritma Naïve Bayes Classifier dan Support Vector Machine Pada Platform X dan YouTube**

Penulis: **Mahyuda Irwanda(\*)**, M Afdal, Rice Novita, Zarnelly

Berdasarkan hasil review dari reviewer bahwa artikel tersebut dinyatakan **DITERIMA** untuk dipublikasikan pada **Volume 7, Nomor 1, June 2025**.

QR Code dibawah ini merupakan penanda keaslian LOA yang dikeluarkan yang akan menuju pada halaman website Daftar LOA pada Jurnal BITS.

Sebagai informasi tambahan, saat ini jurnal **BUILDING OF INFORMATICS, TECHNOLOGY AND SCIENCE (BITS)** telah Re-Akreditasi dan mendapat Peringkat **SINTA 3** berdasarkan SK Kemendikbudristek No. [72/E/KPT/2024](#) tertanggal 1 April 2024 dimulai dari **Volume 5 No 1 (2023)**, hingga **Volume 9 No 4 (2028)**. Sertifikat silahkan diunduh pada link berikut: [[Sertifikat](#)]

Demikian informasi yang kami sampaikan, atas perhatiannya kami ucapkan terimakasih.



Hormat Kami,



**Mesran, M.Kom**  
Journal Manager

Tembusan:

1. Peringgal
2. Author
3. FKPT



Lampiran Surat :

Nomor : Nomor 25/2021

Tanggal : 10 September 2021

## LEMBAR PERNYATAAN

Saya yang bertandatangan di bawah ini:

: Mahyuda Irwanda

: 12150315127

gl. Lahir : Lubuk Gaung/18 juni 2003

Pascasarjana : Sains dan Teknologi

: Sistem Informasi

Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya\*:

ANALISIS SENTIMEN LIGA INDONESIA MENGGUNAKAN ALGORITMA MACHINE

LEARNING PADA PLATFORM X DAN YOUTUBE

Menyatakan dengan sebenar-benarnya bahwa:

1. Penulisan Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya\* dengan judul sebagaimana tersebut diatas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya\* saya ini, saya nyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya\* saya tersebut, maka saya bersedia menerima sanksi peraturan perundang-undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 07 Juli 2025

Yang membuat pernyataan



MAHYUDA IRWANDA

NIM: 12150315127

\* Pilih salah satu sesuai jenis karya tulis





# Analisis Sentimen Masyarakat Terhadap Liga Indonesia Menggunakan Algoritma Naïve Bayes Classifier dan Support Vector Machine Pada Platform X dan YouTube

Mahyuda Irwanda\*, M Afdal, Rice Novita, Zarnelly Zarnelly

Fakultas Sains dan Teknologi, Sistem Informasi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: <sup>1,\*</sup>12150315127@students.uin-suska.ac.id, <sup>2</sup>m.afdal@uin-suska.ac.id, <sup>3</sup>rice.novita@uin-suska.ac.id,

<sup>4</sup>zarnelly@uin-suska.ac.id

Email Penulis Korespondensi: 12150315127@students.uin-suska.ac.id.

Submitted: 06/05/2025; Accepted: 31/05/2025; Published: 01/06/2025

**Abstrak**—Liga Indonesia merupakan ajang kompetisi sepak bola nasional yang menarik banyak perhatian masyarakat. Namun, berbagai permasalahan seperti keputusan wasit yang kontroversial, keributan suporter, dan isu pengaturan skor sering menjadi sorotan. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap Liga Indonesia dengan menggunakan algoritma Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM). Data dikumpulkan dari platform media sosial X (Twitter) sebanyak 2000 tweet dan YouTube sebanyak 2000 komentar dalam rentang waktu Januari 2023 hingga Desember 2024. Setelah melalui tahapan preprocessing seperti cleaning, case folding, tokenizing, stopword removal, dan stemming, data diklasifikasikan menjadi sentimen positif, negatif, dan netral. Hasil penelitian menunjukkan bahwa SVM memiliki akurasi lebih tinggi (99%) dibandingkan NBC (85%) dalam analisis sentimen.

**Kata Kunci:** Sentimen Analisis; Liga Indonesia; Media Sosial; Naïve Bayes; Support Vector Machine

**Abstract**—The Indonesian League is a national football competition that attracts a lot of public attention. However, various problems such as controversial referee decisions, fan riots, and match-fixing issues are often in the spotlight. This study aims to analyze public sentiment towards the Indonesian League using the Naïve Bayes Classifier (NBC) and Support Vector Machine (SVM) algorithms. Data were collected from social media platform X (Twitter) as many as 2000 tweets and YouTube as many as 2000 comments in the period from January 2023 to December 2024. After going through preprocessing stages such as cleaning, case folding, tokenizing, stopword removal, and stemming, the data was classified into positive, negative, and neutral sentiments. The results showed that SVM had a higher accuracy (99%) than NBC (85%) in sentiment analysis.

**Keywords:** Indonesian League; Naïve Bayes; Support Vector Machine; Social Media; Sentiment Analysis

## 1. PENDAHULUAN

Liga Indonesia adalah kompetisi sepak bola nasional yang terdiri dari tiga divisi: Liga 1, Liga 2, dan Liga 3. Semakin tinggi divisi, semakin ketat persaingannya. Tim dengan peringkat terbaik di Liga 1 berhak mewakili Indonesia di AFC Club Competitions 2025/26, sesuai regulasi PSSI dan AFC. Klub yang menempati peringkat pertama dalam klasemen akhir BRI Liga 1 2024/25 akan dinyatakan sebagai juara.

Meskipun menjadi daya tarik utama dan simbol gengsi bagi klub-klub lokal, Liga Indonesia masih menghadapi berbagai masalah yang berdampak pada sektor ekonomi, terutama UMKM. Masalah wasit yang sering mendapat kritik meskipun VAR telah diterapkan menyebabkan ketegangan dalam kompetisi dan memicu protes dari berbagai pihak [1]. Selain itu, insiden kerusuhan suporter, seperti tragedi Stadion Kanjuruhan 2022, semakin merusak kepercayaan publik terhadap penyelenggaraan liga dan berdampak pada keberlangsungan usaha kecil di sekitar stadion. UMKM yang bergantung pada jumlah penonton, seperti pedagang makanan, penjual merchandise, dan penyedia jasa transportasi, mengalami penurunan omzet ketika sentimen terhadap liga memburuk akibat masalah wasit dan keamanan. Sebaliknya, jika kompetisi berjalan lancar dan mendapat respons positif, UMKM memperoleh manfaat ekonomi yang signifikan dari peningkatan jumlah penonton dan peluang kolaborasi dengan klub serta komunitas suporter.

Selain berdampak pada UMKM, permasalahan dalam Liga Indonesia juga memengaruhi pemerintah, terutama dalam hal kebijakan ekonomi dan pengelolaan industri sepak bola. Ketika liga mendapat citra negatif akibat isu seperti pengaturan skor dan ketidakprofesionalan wasit, kepercayaan publik terhadap otoritas sepak bola, termasuk PSSI dan Kemenpora, menurun, sehingga menimbulkan tekanan untuk melakukan reformasi perwasitan, meningkatkan keamanan stadion, serta memperketat regulasi bagi suporter. Pemerintah juga harus menanggung dampak ekonomi dari penurunan jumlah penonton dan investasi di sektor olahraga. Namun, jika liga dikelola dengan baik dan memiliki citra positif, peluang investasi meningkat, wisata olahraga berkembang, dan UMKM yang terkait dengan industri sepak bola dapat tumbuh lebih pesat, yang pada akhirnya berkontribusi pada ekonomi nasional [2]. Oleh karena itu, analisis sentimen masyarakat terhadap Liga Indonesia menjadi penting untuk memahami bagaimana opini publik memengaruhi sektor UMKM, kebijakan pemerintah, serta ekosistem ekonomi secara keseluruhan.

Media sosial seperti X (Twitter) dan YouTube berperan penting dalam membentuk opini publik terhadap sepak bola. X sering digunakan untuk analisis sentimen, khususnya terkait performa tim nasional Indonesia [3]. Sementara itu, komentar di YouTube mencerminkan berbagai pandangan masyarakat, sebagaimana terlihat dalam analisis sentimen Piala Asia U-23 menggunakan algoritma SVM [4]. Oleh karena itu, kedua platform ini menjadi alat yang efektif untuk memahami sentimen publik terhadap sepak bola di Indonesia.



Analisis sentimen digunakan untuk mengidentifikasi opini yang terkandung dalam sebuah teks dan diterapkan pada berbagai jenis teks, seperti ulasan produk, postingan di media sosial, hingga artikel berita [5]. Teknik ini memungkinkan data tidak terstruktur dikonversi menjadi informasi bernilai dan diklasifikasikan sebagai sentimen positif, negatif, atau netral berdasarkan polaritas emosional yang terkandung di dalamnya [6]. Dalam klasifikasi sentimen, algoritma Machine Learning seperti Naive Bayes Classifier (NBC) dan Support Vector Machine (SVM) sering digunakan, di mana SVM bekerja dengan mencari hyperplane optimal untuk memisahkan kelas dalam ruang fitur berdimensi tinggi [7], sementara NBC memanfaatkan Teorema Bayes untuk menghitung probabilitas suatu kelas berdasarkan fitur yang ada, menjadikannya populer dalam klasifikasi teks karena kemudahan implementasi dan efisiensinya[8].

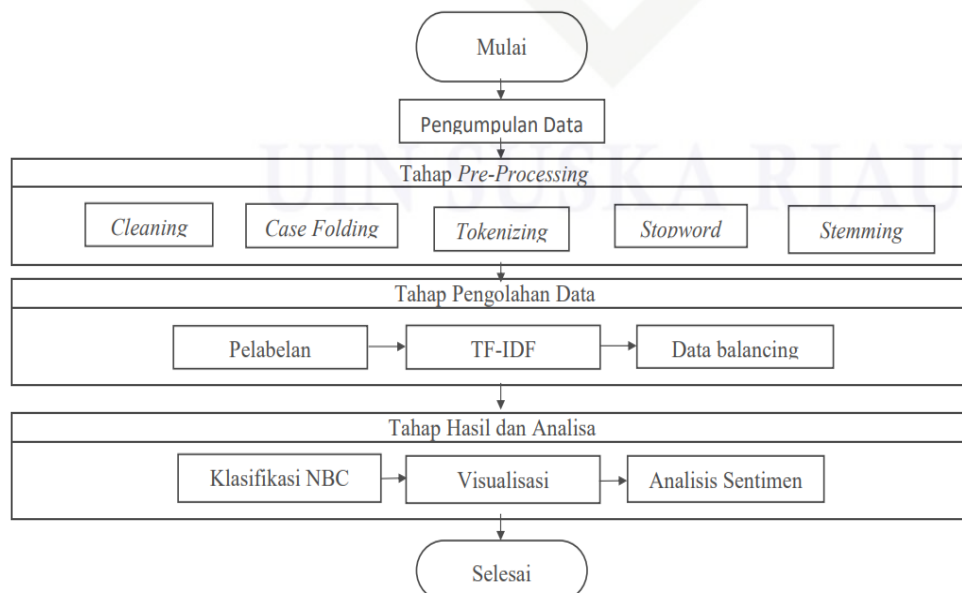
Penelitian sebelumnya yang dilakukan oleh Dalifah (2024) mengenai analisis data sentimen negatif pada opini pengguna twitter terhadap berita sepak bola liga 1 tahun 2022 dengan penerapan support vector machine memperoleh akurasi sebesar 91% [9]. Penelitian yang dilakukan oleh Deni Prasetyo (2021) mengenai analisis sentimen media sosial terhadap bank rakyat indonesia sebagai sponsor resmi liga sepak bola indonesia SVM dengan akurasi 86%. Penelitian yang dilakukan oleh Pratama dkk (2023) mengenai Analisis Sentimen Performa Tim Nasional Sepak Bola Indonesia Era Kepemimpinan Shin Tae-yong pada X Menggunakan Algoritma Naive Bayes mendapatkan akurasi sebesar 90,34%. Penelitian yang dilakukan oleh Alfariis dkk (2024) yang membahas Perbandingan Metode Naive Bayes dan Support Vector Machine pada Analisis Sentimen Aplikasi Mamikos, memperoleh hasil akurasi NBC sebesar 82% Sedangkan SVM sebesar 92%. Penelitian yang dilakukan oleh alfariis yang membahas perbandingan algoritma klasifikasi support vector machine dan naive bayes pada imbalance data, memperoleh hasil peningkatan performa algoritma SVM dari 81,39% menjadi 96,52% dan algoritma NBC dari 82,02% menjadi 94,79%. Hasil ini mengindikasikan bahwa metode SMOTE mampu meningkatkan kinerja algoritma klasifikasi dalam mengatasi permasalahan ketidakseimbangan data dan menunjukkan bahwa, algoritma SVM dapat memprediksi sentimen lebih akurat dan memiliki tingkat kesalahan yang lebih rendah dibandingkan algoritma Naive Bayes [10]. Penelitian ini memiliki perbedaan dari penelitian sebelumnya karena menggabungkan dua platform media sosial, yaitu X (Twitter) dan YouTube, untuk memperoleh data sentimen masyarakat terhadap Liga Indonesia. Selain itu, penelitian ini menggunakan dua algoritma klasifikasi, yaitu Naïve Bayes dan Support Vector Machine (SVM), guna membandingkan akurasi keduanya dalam analisis sentiment.

Tujuan dari penelitian ini adalah untuk menganalisis sentimen masyarakat terhadap Liga Indonesia guna mengetahui persepsi publik, apakah bersifat positif, netral, atau negatif, melalui pendekatan algoritma machine learning. Untuk mencapai tujuan tersebut, penelitian ini membandingkan kinerja dua algoritma klasifikasi, yaitu Naïve Bayes Classifier dan Support Vector Machine, dalam mengklasifikasikan sentimen opini masyarakat yang diambil dari platform X dan YouTube. Hasil dari penelitian ini diharapkan dapat memberikan wawasan mengenai pandangan masyarakat terhadap penyelenggaraan Liga Indonesia.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Untuk mencapai tujuan penelitian, tahapan metodologi yang diterapkan dalam penelitian ini terdapat pada Gambar 1.



Gambar 1. Metode penelitian



## 2.2 Tahap pengumpulan data

Pengumpulan data ulasan dilakukan dengan teknik scraping menggunakan bahasa pemrograman Python, yang mengambil data ulasan pengguna aplikasi X dan YouTube.

## 2.3 Preprocessing

Berikut tahapan text preprocessing yang digunakan pada penelitian ini adalah.

### a. *Cleaning*

Tahap ini bertujuan untuk menghilangkan elemen-elemen yang tidak relevan dari data ulasan guna mengurangi noise dalam proses klasifikasi. Elemen yang dibuang meliputi hashtag, username, tautan URL, simbol, angka, dan tanda baca lainnya. Dengan membersihkan data dari elemen-elemen ini, kualitas data yang diolah dapat meningkat, sehingga analisis dapat menghasilkan hasil yang lebih akurat[11].

### b. *Case folding*

Pada tahap case folding, huruf dalam dokumen diubah ke dalam format standar, biasanya menjadi huruf kecil. Proses ini penting untuk menjaga konsistensi dalam pengolahan teks, sehingga perbedaan antara huruf besar dan kecil tidak mempengaruhi hasil analisis [12].

### c. *Tokenizing*

Tokenizing merupakan metode pengkodean string tertentu berdasarkan penggunaan kata. Hasil dari proses cleaning dan case folding dilanjutkan dengan tahap tokenizing, dimana kalimat review Customer diubah menjadi 15 kata per kata [13].

### d. *Stopword*

Tahap stopword berfungsi untuk menyaring kata-kata penting dari hasil proses sebelumnya. Kata-kata yang tidak memiliki makna dianggap tidak perlu untuk analisis sentimen, sehingga dihapus. Dengan menyaring kata-kata yang relevan, proses analisis menjadi lebih fokus dan efisien[14].

### e. *Stemming*

Pada tahap stemming, kata-kata diproses menjadi bentuk dasar atau dengan menghilangkan imbuhan. Proses ini juga mencakup pengelompokan kata-kata yang memiliki makna dasar yang sama meskipun memiliki bentuk yang berbeda. Hal ini memungkinkan berbagai variasi kata dasar dikenali sebagai satu entitas yang sama, yang akan meningkatkan akurasi dalam analisis teks [15].

## 2.4 Tahap Pengolahan Data

Tahapan ini merupakan proses mengelola, memahami, dan mengkategorikan data teks. Adapun tahapannya sebagai berikut:

### a. *Pelabelan Data*

Proses pelabelan data menjadi positif, negatif dan netral yang dilakukan manual oleh pakar pada data yang sebelumnya telah dilakukan proses pre-processing.

### b. *TF-IDF*

Tahap ini digunakan untuk memberikan bobot pada setiap istilah dalam dokumen. Perhitungan bobot ini dilakukan dengan memanfaatkan modul sklearn dalam bahasa pemrograman Python.

### c. *Data Balancing*

proses data balancing dilakukan dengan menggunakan SMOTE yaitu metode oversampling untuk menyeimbangkan distribusi data antar kelas karena efektivitas analisis sentimen juga dipengaruhi oleh data set yang digunakan, yang dapat mempengaruhi cara penanganannya[16].

## 2.5 Naive Bayes Classifier

NBC adalah algoritma klasifikasi yang sederhana yang menerapkan konsep teorema Bayes dengan anggapan bahwa setiap fitur bersifat independen satu sama lain[17]. Algoritma ini dapat digunakan pada data numerik maupun kategorikal[18]. NBC banyak digunakan dalam berbagai kasus, seperti analisis sentimen, klasifikasi teks, penyaringan spam, dan lain sebagainya [19].

## 2.6 Support Vector Machine

Support Vector Machine(SVM) merupakan satu diantara banyak algoritma yang digunakan untuk klasifikasi dan termasuk dalam kategori supervised learning. Konsep kerja Support Vector Machine yaitu dengan mencari hyperplane atau garis pembatas paling optimal yang berfungsi untuk memisahkan dua kelas [20].

## 2.7 Tahap Hasil dan Analisa

### a. *Klasifikasi Menggunakan Naïve Bayes Classifier dan Support Vector Machine*

Pada tahap klasifikasi ini, dilakukan pengukuran performa model dari dua algoritma, yaitu Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM). NBC adalah metode klasifikasi berbasis probabilitas yang mengasumsikan fitur-fitur data saling independen sehingga sederhana dan cepat dalam pengolahan data teks. SVM adalah algoritma yang bekerja dengan mencari batas pemisah (hyperplane) terbaik yang memaksimalkan jarak antar kelas data, sehingga efektif untuk data kompleks dan. SVM telah banyak diterapkan dan terbukti efektif



dalam analisis sentimen. Langkah ini bertujuan untuk mengevaluasi efektivitas kedua algoritma dalam mengklasifikasikan data berdasarkan sentimen yang terdapat dalam ulasan pengguna.

#### b. Visualisasi Wordcloud

Hasil klasifikasi sentimen kemudian divisualisasikan menggunakan library Wordcloud. Word Cloud adalah representasi data yang menampilkan kumpulan kata-kata penting berdasarkan frekuensi kemunculannya dalam data.

#### c. Analisis Sentimen

Setelah data divisualisasikan data kemudian dianalisis berdasarkan kata yang sering muncul untuk mengidentifikasi opini publik terhadap program sehingga dapat memberikan gambaran yang komprehensif mengenai persepsi masyarakat terhadap Liga Indonesia.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Pengumpulan data

Data yang dikumpulkan pada penelitian ini berasal dari komentar atau tweet masyarakat tentang Liga Indonesia yang terdapat pada platform *X* dan *YouTube*. Proses pengambilan data dilakukan dengan menggunakan teknik pengambilan data otomatis yaitu dengan teknik *Crawling* pada sosial media *X* dan *YouTube*, yang diimplementasikan dengan bahasa pemrograman *Python*. Tweet yang diambil dengan menggunakan hashtag *#LigaIndonesia* dan *#WasitLigaIndonesia*. Total jumlah data yang dikumpulkan adalah 2000 komentar *YouTube* dan 2000 Tweet, dan komentar yang diambil dari 3 video ([perbandingan!!] kualitas penyiaran liga liga asean [siapa yg lebih baik?], kepergian luis milla dan matheus pato menurunkan kualitas liga indonesia!?, optimalkan kualitas liga, mamat yakin bisa tembus timnas, kalo coach justin yang penting etos kerja! ). Total jumlah data yang dikumpulkan adalah 4000 . Adapun data yang diambil ialah data pada tahun 2023 - 2024 (Januari 2023 - Desember 2024).

### 3.2 Proses pre-processing

Setelah data dikumpulkan, dilakukan tahapan pre-processing yang meliputi *cleaning* untuk menghilangkan karakter atau simbol tidak relevan, *case-folding* untuk mengubah teks menjadi huruf kecil, proses *tokenizing* untuk menguraikan teks ke bentuk satuan kata, *stopword removal* untuk menghapus kata-kata minim makna, dan *stemming* untuk mengembalikan kata ke bentuk dasar

**Table 2.** Hasil Pre-Processing Data *X*

| Proses              | Hasil   |
|---------------------|---|
| Data awal           | @SiaranBolaLive Gabung liga indonesia suap wasit suap keeper suap pemain insyaallah 3 bulan 90 goal+ bisa pindah ke liga arab lagi langsung   |
| <i>Cleaning</i>     | SiaranBolaLive Gabung liga indonesia suap wasit suap keeper suap pemain insyaallah bulan goal bisa pindah ke liga arab lagi langsung.   |
| <i>Case folding</i> | siaranbolalive gabung liga indonesia suap wasit suap keeper suap pemain insyaallah bulan goal bisa pindah ke liga arab lagi langsung.   |
| <i>Tokenizing</i>   | ['siaranbolalive', 'gabung', 'liga', 'indonesia', 'suap', 'wasit', 'suap', 'keeper', 'suap', 'pemain', 'insyaallah', 'bulan', 'goal', 'bisa', 'pindah', 'ke', 'liga', 'arab', 'lagi', 'langsung'] |
| <i>Stopword</i>     | ['siaranbolalive', 'gabung', 'liga', 'indonesia', 'suap', 'wasit', 'suap', 'keeper', 'suap', 'pemain', 'insyaallah', 'goal', 'pindah', 'liga', 'arab']  |
| <i>Stemming</i>     | siaranbolalive gabung liga indonesia suap wasit suap keeper suap main insyaallah goal pindah liga arab  |

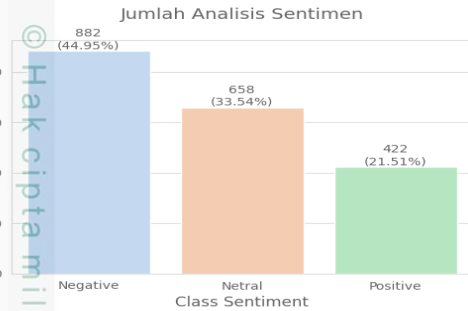
**Tabel 3.** Hasil Pre-Processing Data *YouTube*

| Proses              | Hasil  |
|---------------------|--|
| Data awal           | Masih kaya liga desa desa. Gak lebih. Males nonton karena sering ada keributan mulu.   |
| <i>Cleaning</i>     | Masih kaya liga desa desa Gak lebih Males nonton karena sering ada keributan mulu.   |
| <i>Case folding</i> | masih kaya liga desa desa gak lebih males nonton karena sering ada keributan mulu.   |
| <i>Tokenizing</i>   | ['masih', 'kaya', 'liga', 'desa', 'desa', 'gak', 'lebih', 'males', 'nonton', 'karena', 'sering', 'ada', 'keributan', 'mulu'] |
| <i>Stopword</i>     | ['kaya', 'liga', 'desa', 'desa', 'males', 'nonton', 'keributan', 'mulu']   |
| <i>Stemming</i>     | kaya liga desa desa males nonton ribut mulu.   |

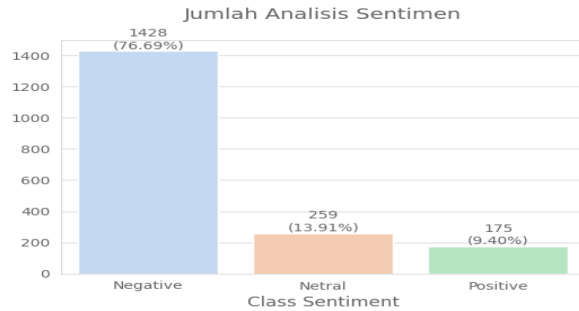
### 3.3 Proses Pelabelan

Dari hasil pelabelan manual oleh pakar sebanyak 2000 data *X* dan 2000 data *YouTube*, didapatkan pada komentar pada Aplikasi *YouTube* memiliki 882 sentimen negatif, 422 sentimen positif, dan 658 sentimen netral. persentase keberhasilan pelabelan setiap kategori dapat dilihat secara rinci pada Gambar 2.





Gambar 2. Persentasi hasil pelabelan data YouTube



Gambar 3. Persentasi hasil pelabelan data X

Pada Gambar 3 menunjukkan bahwa sentimen terhadap aplikasi X didominasi oleh sentimen negatif dengan jumlah 1.428, sementara sentimen positif tercatat sebanyak 175, dan sentimen netral sebanyak 259. .

### 3.4 TF-IDF

Tahap pembobotan TF-IDF untuk mendapatkan bobot dari kata. Hasil yang diperoleh setelah penerapan proses ini terdapat pada gambar 4 dan 5.

| Term      | TF-IDF  | TF  |
|-----------|---------|-----|
| liga      | 0.05221 | 101 |
| main      | 0.03007 | 58  |
| nonton    | 0.02313 | 44  |
| indonesia | 0.02302 | 44  |
| klub      | 0.01764 | 34  |
| timnas    | 0.01676 | 32  |
| el        | 0.01634 | 31  |
| plin      | 0.01592 | 30  |
| sty       | 0.01573 | 30  |
| plan      | 0.01573 | 30  |
| latih     | 0.01495 | 28  |
| bola      | 0.01454 | 28  |
| tonton    | 0.01257 | 24  |
| tiket     | 0.01137 | 22  |
| stadion   | 0.01121 | 21  |

Gambar 4. TF-IDF Data X

| Term      | TF-IDF   | TF |
|-----------|----------|----|
| wasit     | 0.080923 | 78 |
| liga      | 0.07297  | 70 |
| indonesia | 0.064077 | 61 |
| var       | 0.031041 | 29 |
| main      | 0.029324 | 28 |
| kualitas  | 0.028688 | 27 |
| pssi      | 0.02533  | 24 |
| pimpin    | 0.022161 | 21 |
| inggris   | 0.021125 | 20 |
| timnas    | 0.019471 | 18 |
| bola      | 0.018196 | 17 |

Gambar 5. TF-IDF Data YouTube

### 3.5 Data Balancing

Sebelum data diuji dengan algoritma NBC dan SVM data diproses terlebih dahulu menggunakan Random Oversampling untuk menangani ketidakseimbangan kelas pada dataset, hasil data setelah diproses dapat dilihat pada Tabel 4.

Tabel 4. Hasil Penerapan SMOTE Data X

| Label    | Sebelum | Sesudah |
|----------|---------|---------|
| Positif  | 175     | 1428    |
| Negative | 1428    | 1428    |
| Netral   | 259     | 1428    |

Tabel 5. Hasil Penerapan SMOTE Data YouTube

| Label    | Sebelum | Sesudah |
|----------|---------|---------|
| Positif  | 442     | 882     |
| Negative | 882     | 882     |
| Netral   | 658     | 882     |

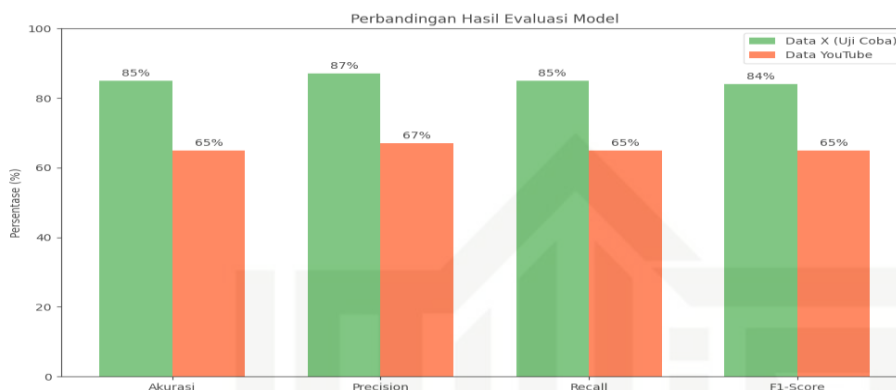
### 3.6 Klasifikasi Algoritma SVM dan NBC

#### 3.6.1 Pembagian Data

Setelah dataset dilakukan proses pre-processing data, pelabelan, dan perhitungan TF-IDF, tahapan selanjutnya ialah melakukan proses klasifikasi menggunakan model SVM dan Random Forest. Sebelum tahapan klasifikasi dilakukan, Data dipisahkan menjadi dua bagian, yaitu data latih dan data uj. Dalam penelitian ini, pembagian data dilakukan dengan proporsi 90:10, yang berarti 90% data dimanfaatkan sebagai data pelatihan untuk pembelajaran model, Sedangkan 10% dari data digunakan untuk data uji dalam mengevaluasi performa model.

### 3.6.2 Hasil Algoritma Naïve Bayes Classifier

Karena Naive Bayes tidak memiliki parameter khusus untuk pengujian, pengujian dilakukan satu kali secara acak oleh sistem. Proses pengujian ini menggunakan Google Colaboratory dan bahasa pemrograman Python. Pengujian ini menghasilkan akurasi sebesar 85%, precision sebesar 87%, recall sebesar 85%, dan f1-score sebesar 84%. Sedangkan data dari YouTube menghasilkan akurasi sebesar 65%, precision sebesar 67%, recall sebesar 65%, dan f1-score sebesar 65%. Detail hasil pengujian tersebut dapat dilihat pada Tabel 4.8.

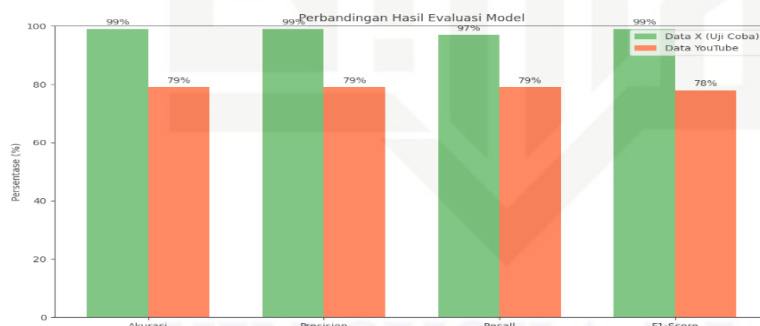


Gambar 4. Hasil Pengujian Algoritma Naïve Bayes Classifier

### 3.6.3 Hasil Algoritma Support Vector Machine

Pada algoritma Support Vector Machine (SVM), evaluasi model dilakukan berdasarkan hasil prediksi tanpa perlu eksplorasi parameter seperti kernel, parameter regulasi (C), dan gamma, karena fokusnya hanya pada pengukuran performa model yang sudah ada. Pengujian pada SVM dilakukan menggunakan parameter default dari library yang digunakan dan dilakukan secara otomatis oleh sistem dengan pembagian data secara acak untuk pelatihan dan pengujian. Pengujian ini menggunakan Google Colaboratory sebagai tools, dan bahasa pemrograman yang digunakan adalah Python.

Pengujian ini menghasilkan akurasi sebesar 99%, precision sebesar 99%, recall sebesar 97%, dan f1-score sebesar 97% untuk data berasal dari aplikasi X, Sedangkan data dari YouTube menghasilkan akurasi sebesar 79%, precision sebesar 79%, recall sebesar 79%, dan f1-score sebesar 78% pada data X. Detail hasil pengujian tersebut dapat dilihat pada Tabel 4.2.



Gambar 5. Hasil Pengujian Algoritma Support Vector Machine

### 3.6.4 Hasil Perbandingan Akurasi Algoritma

Berdasarkan perbandingan kinerja algoritma Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM) pada dua dataset (X dan YT), terlihat bahwa SVM secara konsisten mengungguli NBC dalam semua metrik evaluasi, yaitu Accuracy, Precision, Recall, dan F1-Score. Pada dataset X, NBC mencatat Accuracy 65% sementara SVM mencapai 99%, dan tren serupa terjadi pada metrik lainnya. Sementara itu, pada dataset YT, performa SVM meningkat drastis dengan Accuracy 99% dibandingkan NBC yang hanya 85%. Hal ini menunjukkan bahwa SVM lebih unggul dalam menangani kedua dataset ini, terutama pada dataset YT yang menunjukkan peningkatan signifikan. Meskipun NBC mungkin lebih cepat dan sederhana, SVM terbukti lebih akurat dan dapat diandalkan, sehingga lebih direkomendasikan jika akurasi menjadi prioritas utama.

### 3.6.5 Hasil Perbandingan Sentimen Aplikasi X dan YouTube

Dalam analisis sentimen terhadap Liga Indonesia, perbedaan karakteristik masing-masing platform memengaruhi pola pembentukan serta distribusi opini di kalangan pengguna. Berdasarkan hasil analisis yang dilakukan, ditemukan bahwa di platform YouTube, sentimen negatif terkait Liga Indonesia tercatat sebanyak 1428, sedangkan sentimen

netral mencapai 259, dan sentimen positif berjumlah 175. Sementara itu, di Aplikasi X, sentimen negatif teridentifikasi lebih tinggi, yaitu sebanyak 1.428, dengan sentimen netral sebanyak 259, serta sentimen positif yang tercatat sejumlah 175. Perbedaan distribusi sentimen ini mencerminkan beragamnya cara pengguna dalam mengekspresikan pandangan mereka terhadap Liga Indonesia di kedua platform, yang dapat dipengaruhi oleh faktor seperti format konten, karakteristik audiens, serta tingkat keterlibatan dalam diskusi seputar liga.

### 3.7 Visualisasi Data Sentimen

Visualisasi data bertujuan untuk mengidentifikasi topik yang sering dibahas oleh pengguna X. Dengan memanfaatkan teknik ini, kata yang tersembunyi dalam ribuan komentar dapat diungkap secara lebih jelas dan ringkas sehingga dapat menghasilkan informasi penting terhadap Liga Indonesia. Visualisasi dari kata yang paling sering muncul dari komentar X terhadap Liga Indonesia pada tiap kelas sentimen terdapat pada Gambar 8, 9, dan 10.



Gambar 6. Hasil Visualisasi Sentimen Positif



Gambar 7. Hasil Visualisasi Sentimen Negatif



Gambar 8. Hasil Visualisasi Sentimen Netral

### 3.8 Hasil Analisis Sentimen

Berdasarkan analisis sentimen terhadap 2000 data komentar YouTube dan 2000 tweet di platform X, ditemukan bahwa sebanyak 1.428 data memiliki sentimen negatif, 175 data memiliki sentimen positif, dan 259 data memiliki sentimen netral pada komentar di aplikasi X. Sementara itu, komentar pada YouTube menunjukkan bahwa sebanyak 882 data memiliki sentimen negatif, 422 data memiliki sentimen positif, dan 658 data memiliki sentimen netral. Ketiga kategori ini memberikan gambaran komprehensif tentang bagaimana masyarakat merespons Liga Indonesia.

Pada sentimen positif, hasil wordcloud memperlihatkan dominasi kata-kata seperti "Liga", "Indonesia", "Wasit", dan "Kualitas". Kata-kata ini menunjukkan apresiasi masyarakat terhadap kompetisi Liga Indonesia, terutama dalam aspek peningkatan kualitas permainan dan penyelenggaraan liga. Dukungan terhadap peningkatan infrastruktur stadion, pembinaan pemain muda, serta kualitas siaran pertandingan menjadi faktor utama dalam respons positif ini. Kata "dukung" yang muncul dalam wordcloud mencerminkan semangat masyarakat untuk terus mendukung perkembangan sepak bola nasional.





Pada sentimen negatif, kata-kata yang mendominasi adalah "Wasit", "Tontonan", "Timnas", "Main", dan "Stadion". Kata-kata ini mencerminkan adanya kekhawatiran dan ketidakpuasan masyarakat terhadap aspek tertentu dalam Liga Indonesia. Banyak masyarakat mengkritik kualitas kepemimpinan wasit, transparansi pengelolaan liga, serta keputusan kontroversial dalam pertandingan. Selain itu, kritik juga diarahkan pada kondisi infrastruktur stadion yang masih kurang memadai dan keterbatasan fasilitas yang mendukung kenyamanan penonton. Kata "didik" yang muncul dalam sentimen negatif menunjukkan harapan masyarakat agar pembinaan pemain muda lebih diperhatikan guna meningkatkan kualitas tim nasional Indonesia.

Sementara itu, pada sentimen netral, wordcloud menunjukkan dominasi kata seperti "Wasit", "Liga", "Timnas", "Tonton", dan "Main". Kata-kata ini mengindikasikan bahwa sebagian besar komentar bersifat deskriptif atau hanya menginformasikan terkait pertandingan, hasil skor, serta statistik pemain dan klub. Sentimen netral ini mencerminkan respons yang lebih objektif tanpa ekspresi emosional yang kuat, baik mendukung maupun menentang. Banyak dari komentar ini berasal dari laporan berita, pernyataan resmi dari klub, maupun opini yang hanya mengamati jalannya kompetisi tanpa memberikan kritik atau pujian yang tajam.

Secara keseluruhan, berdasarkan analisis sentimen terhadap komentar publik, ditemukan bahwa proporsi sentimen negatif terhadap Liga Indonesia secara signifikan lebih tinggi dibandingkan sentimen positif atau netral. Kritik utama yang muncul berkaitan dengan kualitas kepemimpinan wasit, transparansi pengelolaan liga, serta infrastruktur stadion yang masih perlu ditingkatkan. Temuan ini menunjukkan bahwa meskipun ada apresiasi terhadap perkembangan Liga Indonesia, masih terdapat berbagai aspek yang perlu diperbaiki untuk meningkatkan kualitas dan kredibilitas kompetisi sepak bola nasional.

## 4. KESIMPULAN

Berdasarkan hasil analisis sentimen terhadap Liga Indonesia, ditemukan bahwa terdapat perbedaan distribusi sentimen antara YouTube dan Aplikasi X, di mana sentimen negatif lebih dominan pada Aplikasi X dengan 1.428 data dibandingkan YouTube yang mencatat 882 sentimen negatif, sementara YouTube memiliki jumlah sentimen netral dan positif yang lebih tinggi. Perbedaan ini menunjukkan bahwa karakteristik platform, seperti format konten, audiens, dan tingkat keterlibatan, memengaruhi ekspresi opini pengguna. Selain itu, hasil pengujian dengan algoritma Naive Bayes Classifier (NBC) dan Support Vector Machine (SVM) menunjukkan bahwa SVM memiliki performa terbaik dalam semua metrik evaluasi. Pada dataset YouTube, SVM mencapai akurasi 79%, lebih tinggi dibandingkan NBC yang hanya 65%, sementara pada dataset X, SVM mencatat akurasi 99% dibandingkan NBC yang hanya 85%. Performa superior SVM ini membuktikan bahwa algoritma tersebut lebih efektif dalam mengklasifikasikan sentimen secara akurat dan konsisten.

## REFERENCES

- [1] D. D. Kuswoyo, H. Pramono, and A. R. RC, "Kontribusi Percaya Diri, Konsentrasi dan Motivasi Terhadap Kinerja Wasit Persatuan Sepak Bola Seluruh Indonesia Provinsi Sumatera Selatan," *Journal of Physical Education and Sports*, vol. 6, no. 3, pp. 241–247, 2017, doi: 10.15294/jpes.v6i3.20587.
- [2] Ajie Wicaksono and Maximianus Agus Prayudi, "Analisis Dampak Penyelenggaraan Fifa World Cup U-17 Pada Sektor Pariwisata Di Indonesia," *EDUTOURISM Journal Of Tourism Research*, vol. 6, no. 01, pp. 90–101, Jun. 2024, doi: 10.53050/ejtr.v6i01.760.
- [3] T. Juniardi and C. A. Sugianto, "Analisis Sentimen Tim Nasional Sepak Bola Indonesia Di Turnamen Piala Dunia U-17 Indonesia Pada Twitter (X) Menggunakan Algoritma Naive Bayes," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3S1, Oct. 2024, doi: 10.23960/jitet.v12i3S1.5188.
- [4] D. Pangestu, M. Malik, and M. R. Pribadi, "Analisis Sentimen Hasil Pertandingan Sepakbola Timnas Indonesia di Piala Asia U-23 pada Platform Youtube menggunakan Algoritma Support Vector Machine (SVM)," *Applied Information Technology and Computer Science (AICOMS)*, vol. 3, no. 1, pp. 38–48, 2024, doi: 10.58466/aicoms.v3i1.1528.
- [5] J. K. Kim, A. Khondker, M. E. Chua, M. Rickard, and A. Lorenzo, "Sentiment analysis of U.S. News & World Report Best Children's Hospital urology rankings: A difference in positivity between the public and academic worlds," *J Pediatr Urol*, vol. 20, pp. S81–S85, Jan. 2024, doi: 10.1016/j.jpuro.2024.06.001.
- [6] O. Alsemaree, A. S. Alam, S. S. Gill, and S. Uhlig, "Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions," *Heliyon*, vol. 10, no. 9, May 2024, doi: 10.1016/j.heliyon.2024.e27863.
- [7] R. Zhang, Y. Li, Y. Gui, D. J. Armaghani, and M. Yari, "A stacked multiple kernel support vector machine for blast induced flyrock prediction," *Geohazard Mechanics*, vol. 2, no. 1, pp. 37–48, Mar. 2024, doi: 10.1016/j.ghm.2024.01.002.
- [8] G. K. Locarso, "Analisis Sentimen Review Aplikasi Pedulilindungi Pada Google Play Store Menggunakan NBC," *Jurnal Teknik Informatika Kaputama (JTik)*, vol. 6, no. 2, 2022, doi: https://doi.org/10.59697/jtik.v6i2.207.
- [9] N. Dalifah, N. Suarna, and W. Prihartono, "Analisis Data Sentimen Negatif Pada Opini Pengguna Twitter Terhadap Berita Sepak Bola Liga 1 Tahun 2022 Dengan Penerapan Support Vector Machine," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, pp. 209–214, Feb. 2024, doi: 10.36040/jati.v8i1.8303.
- [10] C. E. Puspita, O. N. Pratiwi, and E. Sutoyo, "Perbandingan Algoritma Klasifikasi Support Vector Machine Dan Naive Bayes Pada Imbalance Data," *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 8, no. 1, pp. 11–18, Dec. 2021, doi: 10.33330/jurteks.v8i1.1185.
- [11] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 2, p. 406, Apr. 2021, doi: 10.30865/mib.v5i2.2835.

- [12] R. Maulana, A. Voutama, and T. Ridwan, "Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store menggunakan Algoritma NBC," *Jurnal Teknologi Terpadu*, vol. 9, no. 1, pp. 42–48, Jul. 2023, doi: 10.54914/jtt.v9i1.609.
- [13] Alfandi Safira and F. N. Hasan, "Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier," *ZONasi: Jurnal Sistem Informasi*, vol. 5, no. 1, pp. 59–70, Jan. 2023, doi: 10.31849/zn.v5i1.12856.
- [14] G. Radiena and A. Nugroho, "Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi KAI Acces Menggunakan Metode Support Vector Machine," *Jurnal Pendidikan Teknologi Informasi (JUKANTI)*, vol. 6, no. 1, pp. 1–10, Apr. 2023, doi: 10.37792/jukanti.v6i1.836.
- [15] R. Ulgasesa, A. B. P. Negara, and T. Tursina, "Pengaruh Stemming Terhadap Performa Klasifikasi Sentimen Masyarakat Tentang Kebijakan New Normal," *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 10, no. 3, p. 286, Sep. 2022, doi: 10.26418/justin.v10i3.53880.
- [16] P. S. Yadav, R. S. Rao, A. Mishra, and M. Gupta, "Ensemble methods with feature selection and data balancing for improved code smells classification performance," *Eng Appl Artif Intell*, vol. 139, Jan. 2025, doi: 10.1016/j.engappai.2024.109527.
- [17] A. Rahmadeyan and M. Mustakim, "Seleksi Fitur pada Supervised Learning: Klasifikasi Prestasi Belajar Mahasiswa Saat dan Pasca Pandemi COVID-19," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 9, no. 1, pp. 21–32, May 2023, doi: 10.25077/TEKNOSI.v9i1.2023.21-32.
- [18] P. Kamal and S. Ahuja, "An ensemble-based model for prediction of academic performance of students in undergrad professional course," *Journal of Engineering, Design and Technology*, vol. 17, no. 4, pp. 769–781, Aug. 2019, doi: 10.1108/JEDT-11-2018-0204.
- [19] H. Wisnu, M. Afif, and Y. Ruldevyani, "Sentiment analysis on customer satisfaction of digital payment in Indonesia: A comparative study using KNN and Naïve Bayes," *J Phys Conf Ser*, vol. 1444, no. 1, p. 012034, Jan. 2020, doi: 10.1088/1742-6596/1444/1/012034.
- [20] H. C. Husada and A. S. Paramita, "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Teknika*, vol. 10, no. 1, pp. 18–26, Feb. 2021, doi: 10.34148/teknika.v10i1.311.

Halaman ini Diin-  
gini Uidara-  
Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LAMPIRAN A

### BUKTI *SUBMISSION*

[bits] Submission Acknowledgement inbox x

**Dr. Hetty Rohayani** hettyrohayani@gmail.com [via](#) seminar-id.com  
to me ▾

10:46 PM (8 minutes ago) ★

You have a new notification from Building of Informatics, Technology and Science (BITS):

A new article has been submitted to which an editor needs to be assigned.

Link: <https://ejournal.seminar-id.com/index.php/bits/workflow/submission/7294>

Support Journal

---

[Building of Informatics, Technology and Science \(BITS\)](#)

---

[Building of Informatics, Technology and Science \(BITS\)](#)

↩ Reply

➡ Forward



UIN SUSKA RIAU



## LAMPIRAN B

### BUKTI *REVISION*

#### [bits] Editor Decision

2025-05-15 09:58 PM

Mahyuda irwanda irwanda, M. Afdal, Rice Novita, Zarnelly zarnelly:

We have reached a decision regarding your submission to Building of Informatics, Technology and Science (BITS), "Analisis Sentimen Masyarakat Terhadap Liga Indonesia Menggunakan Algoritma Naïve Bayes Classifier Dan Support Vector Machine Pada Platform X dan YouTube".

Our decision is: Revisions Required

Dr. Hetty Rohayani  
Muhammadiyah Jambi University  
Phone 085266157710  
hettyrohayani@gmail.com

← □ ⓘ 🗑️ ✉️ 📧 ⋮ 2 of 2 < > ✎

[bits] REVIEW KE-2: Perbaiki Kembali Manuscript Anda Inbox x ⚙️ 🔗

Oris Krianto Sulaiman, M.Kom [Oris.ks@ft.uisu.ac.id](mailto:Oris.ks@ft.uisu.ac.id) via [seminar-id.com](mailto:seminar-id.com) Tue, May 20, 10:50 PM ★ 😊 ↶ ⋮  
to me ▾

{\$authorName):

Dear Author, setelah kami melakukan pengecekan terhadap artikel yang telah direvisi pada Jurnal **Building of Informatics, Technology and Science (BITS)** dengan judul: "Analisis Sentimen Masyarakat Terhadap Liga Indonesia Menggunakan Algoritma Naive Bayes Classifier Dan Support Vector Machine Pada Platform X dan YouTube", namun masih perlu perbaikan diantaranya:

1. Kajian tentang metode wajib di tambahkan
2. Perbaiki kembali metadata referensi, susunan format IEEE itu : Nama Author, Judul Penelitian, Nama Jurnal, Vol, No, Tahun dan DOI, Wajib title case

Harap perbaiki manuscript anda sesuai dengan komentar reviewer dan nama file harap diberikan urutan revisinya.  
Terimakasih.

Oris Krianto Sulaiman, M.Kom  
Universitas Islam Sumatera Utara, Medan  
[Oris.ks@ft.uisu.ac.id](mailto:Oris.ks@ft.uisu.ac.id)

---

[Building of Informatics, Technology and Science \(BITS\)](#)

↶ Reply ↷ Forward 😊

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LAMPIRAN C

### BUKTI *ACCEPTED*

[bits] Editor Decision ([bits] Keputusan Editor)

✕ 🖨️ 📧

Sat, May 31, 8:25 AM ★ 😊 ↶ ⋮

**Dr. Hetty Rohayani** [hettyrohayani@gmail.com](mailto:hettyrohayani@gmail.com) via [seminar-id.com](#)  
to me, Afdal, Rice, Zarnelly ▾

🗨️ English → Indonesian  
[Show original](#)



Mahyuda irwanda irwanda, M. Afdal, Nasi Novita, Zarnelly zarnelly :

Kami telah mencapai keputusan terkait pengajuan Anda ke Gedung Informatika, Teknologi, dan Sains (BITS), "Analisis Sentimen Masyarakat Terhadap Liga Indonesia Menggunakan Algoritma Naive Bayes Classifier Dan Support Vector Machine Pada Platform X dan YouTube".

Keputusan kami adalah: Menerima Pengajuan

Dr. Hetty Rohayani  
Universitas Muhammadiyah Jambi  
Telepon 085266157710  
[hettyrohayani@gmail.com](mailto:hettyrohayani@gmail.com)

[Gedung Informatika, Teknologi dan Sains \(BITS\)](#)

UIN SUSKA RIAU

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LAMPIRAN D

### BUKTI PELABELAN

#### SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini :

|                  |   |                                   |
|------------------|---|-----------------------------------|
| Nama             | : | Nova Juvita, S.Pd                 |
| Tempat/tgl lahir | : | Siak, 11 November 1994            |
| Pekerjaan        | : | Guru                              |
| Alamat           | : | Sungai selodang Kec.Sungai Mandau |

Menyatakan dengan sesungguhnya bahwa benar memvalidasi untuk pelabelan sentimen yang bersifat Positif, Negatif, dan Netral dalam data tweet X dan *YouTube* Terhadap sentiment masyarakat terhadap Liga Indonesia sebanyak 2.000 data dalam Tugas Akhir dari :

|          |   |  |
|----------|---|--|
| Nama     | : | Mahyuda Irwanda  |
| Nim      | : | 12150315127  |
| Jurusan  | : | Sistem Informasi   |
| Judul TA | : | ANALISIS SENTIMEN MASYARAKAT TERHADAP LIGA INDONESIA<br>MENGUNAKAN ALGORITMA NA'IVE BAYES CLASSIFIER DAN<br>SUPPORT VECTOR MACHINE |

Demikian surat pertanyaan ini saya buat dengan sesungguhnya tanpa paksaan dari pihak manapun. Atas pengertiannya saya ucapkan terimakasih.

Sungai Mandau, Februari 2025

Yang membuat pernyataan



(Nova Juvita, S.Pd)

UIN SUSKA RIAU



## DAFTAR RIWAYAT HIDUP



Mahyuda Irwanda, lahir di Desa Lubuk Gaung, Kecamatan Siak Kecil, Kabupaten Bengkalis, Provinsi Riau pada tanggal 18 Juni 2003. Peneliti merupakan anak bungsu dari dua bersaudara, putra dari pasangan Mulyadi dan Suyati. Pengalaman Pendidikan dimulai dengan menyelesaikan Sekolah Dasar Negeri 03 Benteng Hilir, Kecamatan Mempura, Kabupaten Siak pada tahun 2009-2015. kemudian melanjutkan pendidikan di SMP IT Al-Fath Mempura dan lulus pada tahun 2018. Setelah tamat dari tingkatan sebelumnya, peneliti melanjutkan pendidikan di SMA IT Al-Fath Mempura dan lulus pada tahun 2021. Setelah menyelesaikan pendidikan di bangku sekolah, peneliti melanjutkan pendidikan Strata Satu (S1) di Universitas Islam Negeri Sultan Syarif Kasim Riau, tepatnya Fakultas Sains dan Teknologi, Program Studi Sistem Informasi dan diharapkan dapat menyelesaikan studi pada tahun 2025. Peneliti berhasil menyelesaikan pendidikan S1 dengan judul penelitian “Analisis Sentimen Masyarakat Terhadap Liga Indonesia Menggunakan Algoritma *Naive Bayes* dan *Support Vector Machine* pada Aplikasi X dan YouTube.” Untuk menjalin komunikasi lebih lanjut, peneliti dapat dihubungi melalui *email* 12150315127@students.uin-suska.ac.id.

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.