

Analisis Sentimen dan Topic Modelling Terhadap Tim Nasional Indonesia di Kejuaraan AFF Suzuki Cup 2020 Berdasarkan Opini Pengguna Twitter

Analysis of Sentiment and Topic Modeling of the Indonesian National Team in the 2020 AFF Suzuki Cup Championship Based on Twitter User Opinions

Muhamad Hadiyan Sidik¹, Sri Widiyanesti², Dian Puteri Ramadhani³

¹ Manajemen Bisnis Telekomunikasi dan Informatika, Fakultas Ekonomi & Bisnis, Universitas Telkom, Indonesia, hadiyansidik@student.telkomuniversity.ac.id

² Manajemen Bisnis Telekomunikasi dan Informatika, Fakultas Ekonomi & Bisnis, Universitas Telkom, Indonesia, widiyanesti@telkomuniversity.ac.id

³ Manajemen Bisnis Telekomunikasi dan Informatika, Fakultas Ekonomi & Bisnis, Universitas Telkom, Indonesia, dianpramadhani@telkomuniversity.ac.id

Abstract

Football is the most popular sport in Indonesia. During the Covid-19 pandemic, several football tournaments have been postponed. Most Indonesians do not agree with the postponement. When the AFF Suzuki Cup championship was held, the enthusiasm of the Indonesian people for the tournament was very high. The topic of the national team and the championship is trending on Twitter social media. This study aims to find out how the national team's supporters perceive the national team and to find out the sentiment, the overall topic of conversation, the topic of conversation with positive and negative sentiments of Twitter users towards the Indonesian National Team in the 2020 AFF Suzuki Cup championship. The study uses a naive bayes algorithm to classify. Naive Bayes algorithm is used because it has a high level of accuracy in classifying. The topic modeling research was conducted using Latent Dirichlet Allocation (LDA). This research uses Rstudio and Rapid Miner software.

There are 10,000 reviews collected at the time specified for the study. The data used for the dataset is 1247 tweets with a division of 70% for training data and 30% for test data. The results obtained in the naive Bayes method get a positive sentiment of 62% and 38% have a negative sentiment. These results illustrate that the results of tweets regarding the Indonesian national team in the AFFC Suzuki Cup 2020 event are positive.

Keywords-consumer perception, naive bayes, sentiment analysis, topic modeling, twitter

Abstrak

Sepak bola adalah olah raga yang paling digemari oleh masyarakat Indonesia. Selama pandemi Covid-19, beberapa turnamen sepak bola mengalami penundaan. Sebagian besar masyarakat Indonesia tidak setuju dengan penundaan tersebut. Saat kejuaraan AFF Suzuki Cup dilaksanakan, antusiasme masyarakat Indonesia terhadap turnamen tersebut sangat tinggi. Topik mengenai tim nasional dan kejuaraan menjadi trending di media sosial Twitter. Penelitian ini bertujuan mengetahui bagaimana persepsi suporter tim nasional terhadap tim nasional dan mengetahui sentimen, topik pembicaraan keseluruhan, topik pembicaraan bersentimen positif dan negatif pengguna Twitter terhadap Tim Nasional Indonesia di kejuaraan AFF Suzuki Cup 2020. Penelitian menggunakan algoritma naive bayes untuk melakukan klasifikasi. Algoritma naive bayes digunakan dikarenakan mempunyai tingkat akurasi yang tinggi dalam melakukan klasifikasi. Penelitian topic modelling dilakukan dengan menggunakan Latent Dirichlet Allocation (LDA). Penelitian ini menggunakan software Rstudio dan Rapid Miner. Terdapat 10.000 ulasan yang terkumpul pada waktu yang telah ditentukan untuk dilakukan penelitian. Data yang digunakan untuk dataset serbanyak 1247 tweet dengan pembagian 70% untuk data latih dan 30% untuk data uji. Hasil yang diperoleh pada metode naive bayes mendapatkan hasil sentimen positif sebesar 62% dan 38% memiliki sentimen negatif. Hasil tersebut menggambarkan bahwa hasil tweet yang mengenai tim nasional Indonesia di ajang AFFC Suzuki Cup 2020 adalah positif.

Kata kunci-analisis sentimen, naive bayes, persepsi konsumen, topic modeling, twitter.

I. PENDAHULUAN

A. Latar Belakang

Sepak bola merupakan olah raga yang paling digemari oleh masyarakat Indonesia (Pradiksa dan Prianto, 2022). Sebanyak 90,8% penduduk Indonesia mengetahui sepak bola dan 46,7% dari penduduk yang mengetahui tersebut menyukai dan memiliki minat dalam sepak bola (Pradiksa dan Prianto, 2022). Indonesia memiliki beberapa kelompok suporter sepakbola, baik itu suporter klub lokal (Virantika, 2021) hingga suporter klub luar negeri. Suporter sepak bola di Indonesia dikenal sebagai suporter bola yang fanatik (Bleacher Report, 2019). Suporter sepak bola di Indonesia memiliki jumlah yang sangat banyak bahkan beberapa kelompok suporter di Indonesia

masuk kedalam sepuluh besar kelompok suporter dengan rata rata keterisian kursi stadion terbanyak di asia tenggara pada tahun 2019 (Fatoni,2019).

Pandemi Covid-19 terjadi pada tahun 2020 sampai 2022. Turnamen sepak bola beberapa ada yang ditunda salah satunya adalah liga Indonesia (Rahayu, 2021) dan AFF Suzuki Cup yang ditunda hingga akhir tahun 2021 (Mukti, 2021). Penundaan turnamen ini menimbulkan beberapa respon dari masyarakat, mayoritas masyarakat tidak setuju dengan penundaan turnamen sepak bola selama masa pandemi dan setuju agar turnamen sepak bola tetap digelar tanpa kehadiran suporter di stadion (Pradiksa dan Prianto, 2022).

AFF Suzuki Cup dilaksanakan pada akhir tahun 2021 dan tim nasional sepak bola Indonesia mengikuti kejuaraan AFF Suzuki Cup. Tim nasional Indonesia berhasil menjadi *runner up* di kejuaraan ini (PSSI, 2021). Pada saat turnamen dilaksanakan, topik mengenai kejuaraan dan tim nasional Indonesia sering menjadi trending topik di media sosial Twitter. Gambar 1 menunjukkan trending di media sosial Twitter selama kejuaraan AFF Suzuki Cup berlangsung.



Gambar 1 Trend Populer Twitter (Twitter)

Wicaksono dalam (Rijal et.al, 2021) keberadaan suporter sangat penting bagi sebuah klub dikarenakan mereka dapat memberikan nuansa yang mendukung para pemain yang berada di lapangan dan mereka juga menjadi sumber pemasukan bagi sebuah klub salah satunya melalui penjualan tiket. Keberadaan suporter sebagai pelanggan sangatlah penting dalam industri sepakbola (Sumarno dan Irianto, 2019).

PSSI dapat mengetahui persepsi suporter serta meningkatkan kualitas Tim Nasional Indonesia dengan cara memanfaatkan opini suporter terhadap tim nasional Indonesia di media sosial *Twitter*. Media sosial *twitter* adalah salah satu sosial media paling populer di dunia dengan pertumbuhan jumlah pengguna yang terus meningkat sehingga dapat dijadikan sarana untuk menuliskan kritik, saran dan juga opini (Rozi et al., 2018).

Opini suporter yang terdapat dalam media sosial *Twitter* mengandung banyak dukungan, kritik serta saran bagi tim nasional Indonesia. Opini tersebut dapat dimanfaatkan untuk dijadikan masukan bagi tim sehingga dibutuhkan metode yang dapat mengolah data berupa opini tersebut dengan cepat dan otomatis mengklasifikasikan data tersebut kedalam kategori positif atau negatif.

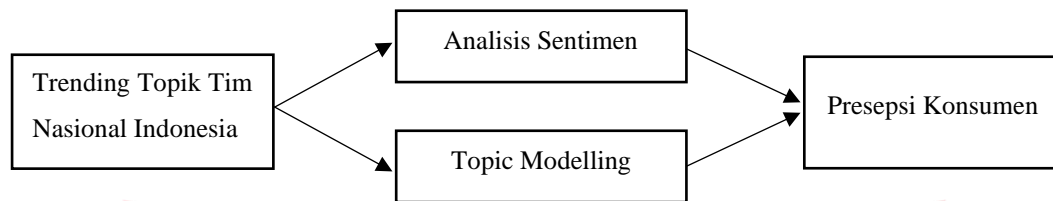
Menurut Ipmawati, J. et al. (2017) analisis sentimen adalah sebuah proses yang bertujuan untuk mendapatkan informasi dari sebuah data tekstual sehingga dapat mengetahui gambaran umum/presepsi masyarakat terhadap suatu topik/objek berupa produk atau layanan. Medhat et al. (Hudaya et al., 2019) berpendapat bahwa analisis sentimen mengklasifikasikan pendapat/emosi menjadi positif atau negatif dari sebuah kalimat.

Nugroho (2014) berpendapat bahwa *Naive Bayes* merupakan sebuah algoritma pengklasifikasian yang digunakan untuk memprediksi kemungkinan keanggotaan suatu kelas. *Topic modelling* merupakan salah satu teknik *Natural Language Processing* yang bertujuan untuk menganalisis teks (aggarwal dan Zhai dalam Naury et. al, 2021).

B. Perumusan Masalah

Metode yang digunakan adalah metode *naive bayes* dan *topic modelling*. *Naive bayes* digunakan untuk mengklasifikasi sentimen dari tweet yang digunakan sebagai data. Selain *naive bayes* ada beberapa metode lain dalam melakukan klasifikasi seperti algoritma C4.5, *K-means*, *Priori*, *Naive Bayes* dan *Support Vector Machines*. Metode yang digunakan untuk melakukan pemodelan topik ada beberapa macam seperti *Latent Semantic Analysis (LSA)*, *Probabilistic Latent Sematic Analysis (pLSA)* dan *Latent Dirichlet Allocation (LDA)* (Nurlayli dan Nasichuddin, 2019). Menurut Campbell et.al dalam (Alfanzar, 2019), salah satu metode yang paling banyak digunakan saat ini untuk melakukan *topic modelling* dan topik analisis adalah *Latent Dirichlet Allocation (LDA)*,

dikarenakan dapat menghasilkan daftar topik yang diberi bobot sehingga dapat digunakan untuk merangkum, mengelompokan, menghubungkan maupun memproses data yang sangat besar.



Gambar 3 Kerangka Pemikiran

Penelitian ini berlandaskan teori mengenai persepsi konsumen yang merupakan rangsangan yang dirasakan dan diterima oleh konsumen sehingga konsumen memproses, mengartikan dan mentafsirkan informasi untuk mendapatkan gambaran dari suatu hal yang memiliki arti bagi pelanggan (Panjaitan, 2018).

Analisis sentimen adalah proses yang mempunyai tujuan untuk mendapatkan informasi dari data tekstual sehingga dapat mengetahui gambaran umum/presepsi masyarakat terhadap suatu topik/objek (Ipmawati et al., 2017). Analisis sentimen mengklasifikasikan pendapat/emosi menjadi positif atau negatif dari data yang diteliti (Medhat et al. dalam Hudaya et al., 2019)

Topic modelling ini merupakan algoritma yang dapat mengidentifikasi pola dalam serangkaian kata di dalam sekumpulan dokumen (Jacobi et. al dalam Naury et. al, 2021). Menurut Campbell et.al dalam (Alfanzar, 2019) Topic Modelling dapat menghasilkan daftar topik yang diberi bobot sehingga dapat digunakan untuk merangkum, mengelompokan, menghubungkan maupun memproses data yang sangat besar.

II. METODOLOGI PENELITIAN

A. Jenis Penelitian

Pada penelitian ini, penulis menggunakan metode kuantitatif. Metode ini dapat secara akurat mengukur opini, sikap, pengetahuan serta perilaku (Indrawati, 2015). Metode ini dapat menghasilkan analisis yang lebih akurat dikarenakan angka statistik dimanfaatkan dalam proses membuat kesimpulan (Mardianie, 2018). Penelitian deskriptif mempunyai tujuan untuk menjelaskan gambaran lengkap dari suatu fenomena yang ada (Cahyono, 2016). Penelitian yang penulis lakukan bertujuan untuk menganalisis sentimen serta topic modelling dari opini supporter tim nasional Indonesia di sosial media Twitter. Penulis tidak mengintervensi data yang diambil dari Twitter. Unit analisis dalam penelitian adalah individual dimana penulis menganalisis opini dari supporter terhadap tim nasional Indonesia. Waktu pelaksanaan dalam penelitian ini adalah cross sectional. Penelitian ini mencari hubungan antar variabel yang diidentifikasi di waktu yang sama. Subyek penelitian pun hanya diobservasi sekali (Widyaningsih et. Al, 2019)

Tabel 3. Variabel Penelitian

No	Variabel	Definisi Variabel
1	Tweet/kicauan	Postingan berupa tulisan di sosial media Twitter
2	Data terlabel	Data yang sudah memiliki nilai berdasarkan kamus kata positif dan negatif
3	Klasifikasi	Tulisan yang sudah memiliki kategori sentimen

Sumber: Olahan Penulis

B. Operasional Variabel

Variabel yang digunakan dalam penelitian yang dilakukan akan disebutkan serta dijelaskan dalam tabel 3.

Variabel	Sub Variabel	Definisi Operasional	Indikator
Sentimen		Respon yang berbeda beda terhadap suatu gejala atau isu (Syahril Iskandar, M., 2020)	
	Positif	Data yang memiliki sentimen positif	Kata bernilai positif

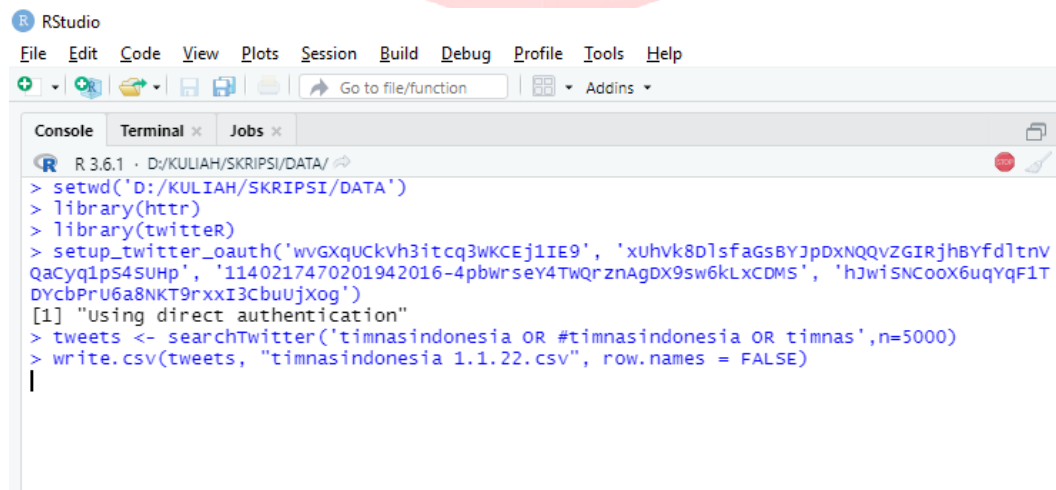
Negatif	Data yang memiliki sentimen Kata bernilai negatif negatif
Topik Pembicaraan	Suatu isu yang menjadi pokok pembicaraan dalam sebuah kalimat (Silaswati, D., 2018).

C. Tahapan Penelitian

Penelitian ini dimulai dengan menentukan latar belakang. Selanjutnya dilakukan identifikasi masalah, penentuan tujuan, studi literatur, mengumpulkan data, *preprocessing* data, analisis dan mengambil kesimpulan dari penelitian yang dilakukan.

D. Pengumpulan Data

Data dikumpulkan dari media sosial *Twitter*. Data yang diambil merupakan *tweet* (*tweet*, *reply*, *retweet*) berbahasa indonesia yang mengandung kata kunci yang telah ditentukan yaitu “timnasindonesia”, “#timnasindonesia” dan “timnas”. Pengambilan data dilakukan dengan menggunakan *software* Rstudio yang dilakukan dalam rentang waktu 22 Desember 2021 - 1 Januari 2022.



```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Console Terminal x Jobs x
R 3.6.1 · D:/KULIAH/SKRIPSI/DATA/
> setwd('D:/KULIAH/SKRIPSI/DATA')
> library(httr)
> library(twitter)
> setup_twitter_oauth('wvGxquckvh3itcq3WkCEj1IE9', 'xuhvk8D1sfaGSBYjpdXNQQVZGIRjhBYfd1tnv
Qacyq1pS4SUHp', '1140217470201942016-4pbwrsey4TWqrznAgDX9sw6kLXCDMS', 'hJwiSNCoox6uqYqF1T
DYcbPrU6a8NKT9rxxI3Cbuujxog')
[1] "using direct authentication"
> tweets <- searchTwitter('timnasindonesia OR #timnasindonesia OR timnas',n=5000)
> write.csv(tweets, "timnasindonesia 1.1.22.csv", row.names = FALSE)
|

```

Gambar 4. Pengambilan Data Dengan menggunakan R Studio (Software R Studio)

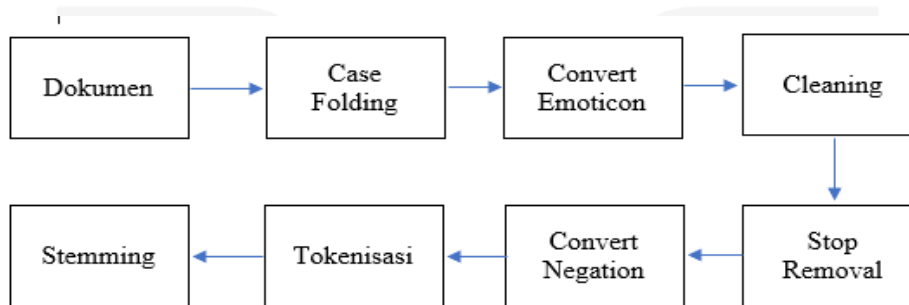
E. Teknik Analisis Data

Pengumpulan data dilakukan dengan cara mengcrawling data dari sosial media *Twitter*. Data yang diambil sebanyak 10.000 *Tweet*. Data yang dihasilkan dari proses crawling memiliki format *CSV*. Data tersebut memiliki beberapa atribut seperti *username*, *tweet* dan atribut lainnya.

Gambar 5. Hasil Pengumpulan Data Berupa Format CSV

F. Preprocessing Data

Data yang didapatkan dari proses crawling merupakan data mentah serta mengandung banyak sekali karakter yang tidak dapat. Dalam tahap *preprocessing* data, dilakukan beberapa tahapan proses agar data menjadi lebih terstruktur sehingga dapat dianalisis. Data yang akan digunakan dalam penelitian ini adalah *username* dan *tweet* berbahasa Indonesia. Berikut proses *preprocessing* data menurut Rustiana dan Rahayu (2017):



Gambar 6. Tahapan *Preprocessing* Data (Rustiana dan Rahayu, 2017)

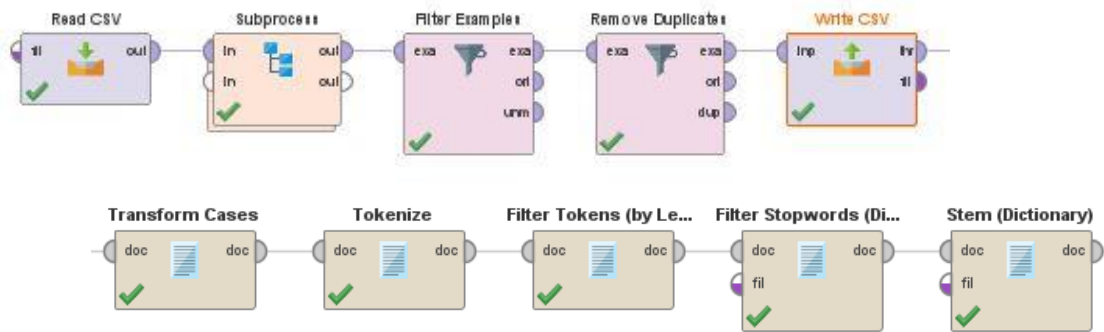
Dari gambar diatas terdapat 7 tahapan preprocessing data. Berikut penjelasan dari tahap tersebut:

1. *Case Folding* merupakan tahapanmenubah huruf kapital diubah menjadi huruf kecil
2. *Convert Emoticon* merupakan tahapan dimana emoticon diubah kedalam string yang sesuai
3. *Cleaning* merupakan tahapan menghapus komponen yang tidak berpengaruh terhadap nilai sentimen
4. *Stop Removal* merupakan tahapan menghapus kata yang tidak berhubungan dengan penelitian
5. *Convert Negation* adalah proses perubahan nilai sentimen kata kata yang bersifat negasi
6. *Tokenisasi* merupakan proses memotong tanda memisah kata
7. *Stemming* merupakan pengubahan kata berimbuhan menjadi kata dasar Rustiana dan Rahayu (2017).

Preprocessing data dilakukan dengan menggunakan software Rapidminer. Preprocessing menggunakan beberapa operator yang tersedia di software Rapidminer, diantaranya:

1. Read CSV untuk membaca file dalam bentuk CSV
2. Subprocess berfungsi untuk menggabungkan operator
3. Filter Examples berfungsi untuk mengfilter data
4. Remove duplicate berfungsi untuk menghapus data yang terduplikat
5. Transform Cases yang berfungsi untuk mengubah karakter kapital menjadi karakter kecil
6. Tokenize yang berfungsi untuk memecahkan kalimat menjadi kata
7. Filter tokens berfungsi untuk membatasi jumlah karakter dalam satu kata.
8. Filter Stopwords untuk menghapus kata yang tidak ada di kamus stopwords.

9. Stem yang berfungsi agar kata yang memiliki imbuhan berubah bentuk menjadi kata dasar.
10. Write CSV berfungsi untuk menyimpan data hasil preprocessing dalam format CSV



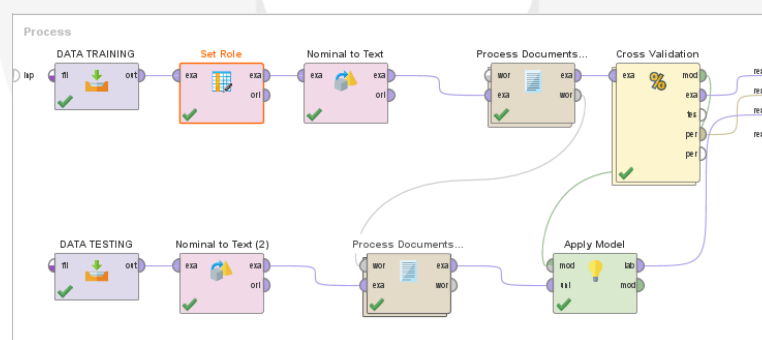
Gambar 7. Preprocessing dalam Rapidminer (Rapidminer)

G. Analisis Sentimen

Medhat et al. (Hudaya et al., 2019:667) berpendapat bahwa analisis sentimen mengklasifikasikan pendapat menjadi positif atau negatif. Mahardika et al. (2018:410) juga menambahkan bahwa analisis sentimen bisa mengklasifikasikan pendapat/emosi menjadi netral. Dalam penelitian ini penulis akan melakukan analisis sentimen yang bertujuan untuk mengelompokkan tweet berupa opini terhadap tim nasional Indonesia kedalam kelompok sentimen positif dan negatif. Data dibagi menjadi dua bagian yaitu data training dan data testing dengan presentase 70% dan 30% dari total data. Data training diberi dua label yaitu positif dan negatif (Alamsyah et. Al, 2018)

Klasifikasi data testing dilakukan dengan menggunakan software Rapidminer. Terdapat beberapa operator dalam proses klasifikasi data testing sebagai berikut:

1. *Read CSV* berfungsi untuk membaca file data dalam bentuk *CSV*
2. *Set Role* berfungsi untuk mengubah peran atribut
3. *Nominal to Text* berfungsi agar atribut *nominal* berubah menjadi atribut *text*
4. *Process Document from Data* berfungsi untuk menyimpan operator *preprocessing*
5. *Cross Validation* berfungsi untuk melakukan validasi silang untuk menilai performa model
6. *Apply Model* berfungsi untuk menerapkan model pada *example set*.



Gambar 8. Proses klasifikasi data menggunakan Rapidminer (Rapid Miner)

H. Confusion Matrix

Menurut Hozairi, Anwari & Alim (2021) Confusion Matrix atau error matrix merupakan sebuah metode yang dapat digunakan untuk mengukur performa/akurasi klasifikasi yang memiliki 4 representasi hasil proses klasifikasi yaitu TP (*True Positif*), TN (*True Negative*), FP (*False Positif*) dan FN (*False Negatif*)

Confusion Matrix ini memiliki tiga komponen yaitu akurasi, presisi dan recall. Menurut Azhari, Situmorang & Rosnelly (2021) presisi merupakan nilai ketepatan sistem dalam menunjukkan data kelas negatif atau kelas positif, sedangkan recall adalah nilai yang menunjukkan tingkat keberhasilan untuk mengetahui kembali secara benar data

kelas positif maupun negatif, sedangkan untuk akurasi adalah nilai rasio data yang benar terdeteksi dalam data pengujian.

Tabel 4. Confusion Matrix

Classification	Predicted Class	
	True	False
Actual: True	TP	FN
Actual: False	FP	TN

Sumber: Hozairi, Anwari & Alim (2021)

$$\begin{aligned}
 \text{Akurasi} &= \frac{\text{True Positif} + \text{True Negatif}}{\text{True Positif} + \text{True Negatif} + \text{False Positif} + \text{False Negatif}} & (3.1) \\
 \text{Precision} &= \frac{\text{True Positif}}{\text{True Positif} + \text{False Positif}} & (3.2) \\
 \text{Recall} &= \frac{\text{True Positif}}{\text{True Positif} + \text{False Negatif}} & (3.3)
 \end{aligned}$$

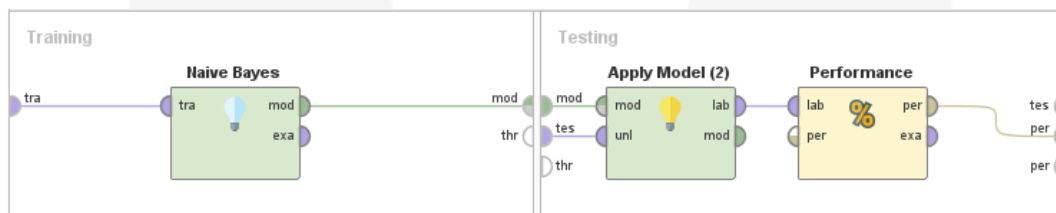
Gambar 6. Performas Model Klasifikasi (Hozairi, Anwari & Alim, 2021)

Evaluasi performansi dilakukan dengan menggunakan tiga operator sebagai berikut:

Naive Bayes berfungsi untuk menghasilkan model klasifikasi naive bayes

Apply Model berfungsi agar model dapat diterapkan pada dataset

Performance berfungsi untuk melakukan evaluasi terhadap kinerja model yang diterapkan



Gambar 9. Proses Evaluasi Performansi (Rapidminer)

a. Topic Modelling

Topic modelling ini merupakan algoritma yang dapat mengidentifikasi pola dalam serangkaian kata di dalam sekumpulan dokumen (Jacobi et. al dalam Naury et. al,2021). Dalam melakukan *topic modelling* penulis melakukan *topic modelling* menggunakan *software Rstudio* dan menggunakan atribut *tweet*.

III. HASIL DAN PEMBAHASAN

A. Karakteristik Data

Data yang digunakan dalam proses penelitian ini adalah tweets berbahasa Indonesia yang mengandung kata kunci “timnasindonesia”, “#timnasindonesia” dan “timnas”. Data yang didapatkan dari proses crawling dengan menggunakan *software Rstudio* mengandung beberapa data yang tidak diperlukan seperti *URL*, *username* dan simbol. Data diambil dalam rentang waktu 23 Desember 2021 – 1 Januari 2022. Jumlah data sebelum *preprocessing* data berjumlah 10.000 tweets. Setelah *preprocessing*, jumlah data berkurang menjadi 1.247 *tweets*.

```

text,"favorited","favoriteCount","replyToSN","created","truncated","replyToSID","id","replyToUID","statusSource","
@tokopedia Dukung timnas indonesia#OPPOA9520.378,FALSE,0,"pejuang_5a",2022-01-02 11:23:30,FALSE,"14732745057
@tokopedia Dukung timnas indonesia#OPPOA9520.353,FALSE,0,"pejuang_5a",2022-01-02 11:23:30,FALSE,"14732745057
@tokopedia Dukung timnas indonesia#OPPOA9520.374,FALSE,0,"pejuang_5a",2022-01-02 11:23:30,FALSE,"14732745057
@tokopedia Dukung timnas indonesia#OPPOA9520.361,FALSE,0,"pejuang_5a",2022-01-02 11:23:30,FALSE,"14732745057
@tokopedia Dukung timnas indonesia#OPPOA9520.37,FALSE,0,"pejuang_5a",2022-01-02 11:23:30,FALSE,"14732745057
@tokopedia Dukung timnas indonesia#OPPOA9520.372,FALSE,0,"pejuang_5a",2022-01-02 11:23:30,FALSE,"14732745057
RT @pandji: Pulang ke Indonesia dgn kepala tegak, timnas.,FALSE,0,NA,2022-01-02 11:23:20,FALSE,NA,"1477601399412
RT @FaktaSepakbola: Sejumlah pemain timnas Indonesia pada squad AFF, yang berkarir di luar negeri (abroad):• Syah
RT @jokowi: Selamat kepada Thailand yang tampil menjadi juara Piala AFF 2020 malam ini. Bagi Timnas Indonesia, kali
RT @TeleskopID: Tapi apapun itu, gue bangga dengan tim ini. Gue juga bangga punya Kapten tim bernama Asnawi Mar
RT @detikcom: Perjuangan Timnas Indonesia mencapai Final Piala AFF 2020 diapresiasi netizen. Dengan skuad rata-rai
RT @chelseaindo: Terima kasih telah memberikan usaha semaksimal mungkin untuk kami semua. Perjuangan kalian s
RT @husainabdullah1: Pengamat bola Malaysia, akhirnya akui kehebatan bintang bintang sepakbola Timnas Indonesia
RT @psi_id: Terima kasih timnas!<U+0001F1EE><U+0001F1E9><U+0001F339><U+270A><U+0001F3FD>,FALSE,0,NA,2022-1
RT @KompasTV: Timnas Indonesia Piala AFF 2020 Disambut Mochamad Iriawan https://t.co/HaZWQ4DAKk https://t.co
Td gue turun ke lobby buat ambil gofood, ga lama timnas indo datang utk karantina di hotel ini. Kurang beruntung ya..
RT @garagarabola : 1. <U+0001F1EE><U+0001F1E9> Indonesia sukses total menciptakan 20 gol. Terproduktif2. <U+000:
RT @panditfootball: Kawal terus perjalanan Timnas Indonesia <U+0001F1EE><U+0001F1E9> #SelamanyaGaruda https://
RT @davidnurbianto: Jadi juara group diatas Vietnam. Mempermalukan Malaysia. Goal cantik mendunia. Big drama la
RT @erickthohir: Leg demi leg telah dilewati dengan determinasi tinggi. Selamat atas prestasi Timnas di Piala AFF 2021
RT @ShopeeID: Terima kasih Timnas Indonesia yang telah memberikan permainan terbaik di Piala AFF 2021. Kerja ker
RT @TeleskopID: Taai apapun itu. eue bangga dengan tim ini. Gue juga bangga punya Kapten tim bernama Asnawi Mar
    
```

Gambar 10. Data sebelum dilakukan preprocessing (data yang sudah diolah)

B. Analisis Sentimen

Preprocessing data dilakukan agar data yang tidak dibutuhkan dapat dihilangkan sehingga kualitas data menjadi lebih baik. Data kemudian dibagi menjadi data testing dan data training. Data dibagi dengan perbandingan 30% (data testing) dan 70% (data training). Data kemudian diklasifikasikan ke dalam sentimen positif dan negatif. Algoritma Naive Bayes dimanfaatkan dalam proses pengklasifikasian data testing. Data yang telah diklasifikasikan kemudian dilakukan evaluasi performansi.

Jumlah data setelah dilakukan preprocessing mengalami penyusutan menjadi berjumlah 1247 tweets. Terdapat 8753 tweets yang dihapus. Data yang sudah melalui preprocessing data selanjutnya dilakukan klasifikasi data. Data dibagi menjadi dua yaitu data training (70%) dan data testing (30%). Data training berjumlah 874 data, sedangkan data testing berjumlah 373 data. Data training diberi label sentimen terlebih dahulu, sedangkan data testing tidak diberi label terlebih dahulu.

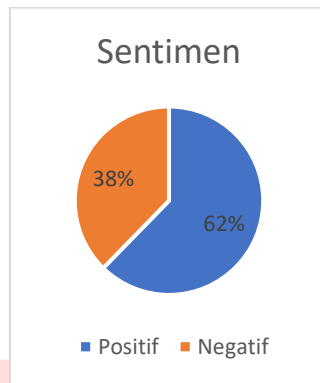
Akurasi dari model yang diterapkan terhadap dataset sebesar 64,41%. Presentase presisi dari keseluruhan kelas yang diprediksi positif sebesar 83,29% sedangkan untuk keseluruhan kelas negatif sebesar 48,85%. Presentase recall dari kelas yang diprediksi positif sebesar 57,32% sedangkan untuk kelas yang diprediksi negatif sebesar 78%

accuracy: 64.41% +/- 6.35% (micro average: 64.42%)

	true positif	true negatif	class precision
pred. positif	329	66	83.29%
pred. negatif	245	234	48.85%
class recall	57.32%	78.00%	

Gambar 11. Hasil evaluasi performansi (Rapidminer)

Klasifikasi dari 1247 tweets yang sudah dilakukan preprocessing dan klasifikasi terdapat 777 tweets yang memiliki nilai sentimen positif dan juga 470 tweets yang memiliki nilai sentimen negatif. Terdapat 62% dari total tweet bersentimen positif dan 38% bersentimen negatif. Tweets dengan sentimen negatif lebih sedikit dari sentimen positif.

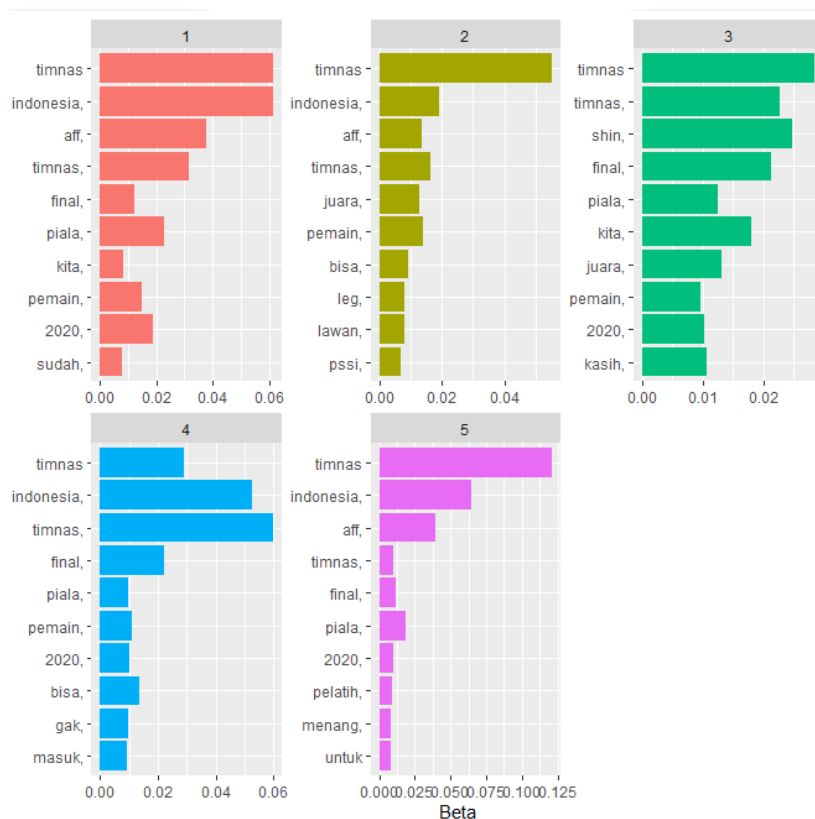


Gambar 12. Hasil Sentimen (Data yang sudah diolah)

C. Topic Modelling

Data yang telah dianalisis sentimen selanjutnya dilakukan pengolahan kembali agar mengetahui topik apa saja yang paling dibahas supporter tim nasional Indonesia di media sosial Twitter. Pengolahan data tersebut menggunakan aplikasi Rstudio dengan memanfaatkan metode LDA.

Metode LDA memiliki tingkat keefektifan tertinggi dalam menemukan topik yang paling sering dibicarakan didalam sebuah dokumen menggunakan Rstudio dengan jumlah kelompok topik yang ditentukan terlebih dahulu sebelum dilakukan proses pemodelan (Salsabila & Trianasari, 2021). Jumlah kelompok topik yang paling optimal dan paling sederhana untuk dilakukan diskusi adalah 5 topik (Masrury, Fannisa & Alamsyah, 2019). Berikut hasil visualisasi topik modelling terhadap 1.247 data yang digunakan dalam pemodelan topik



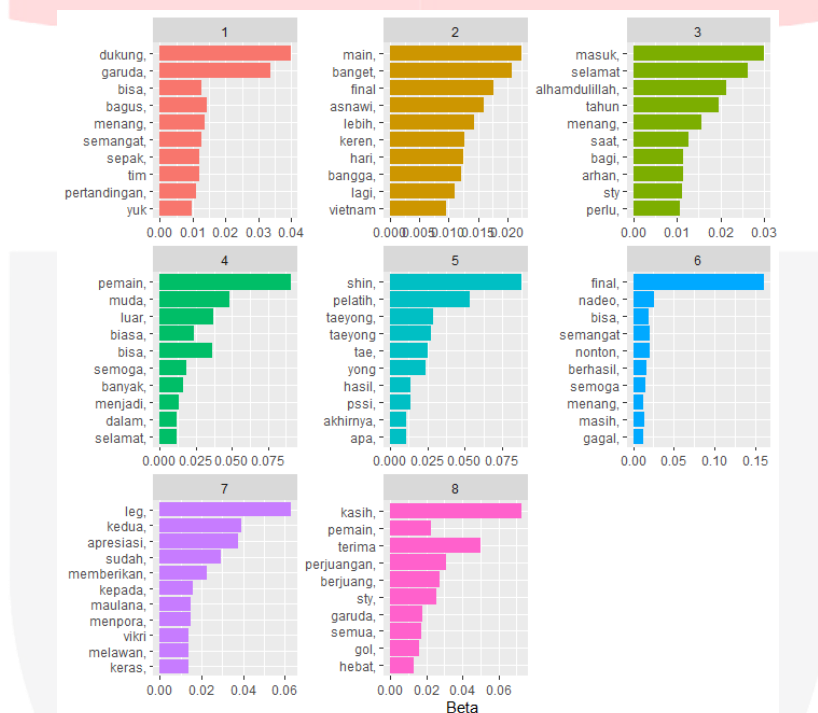
Gambar 14. Visualisasi Hasil Topic Modeling (Data yang telah diolah)

Gambar 14 merupakan visualisasi hasil topic modelling terhadap data yang digunakan. Terdapat lima kelompok topik dalam visualisasi tersebut. Setiap topik mengandung 10 kata dimana setiap kata memiliki nilai beta. Semakin besar nilai beta, semakin sering kata tersebut muncul dalam suatu topik. Topik pertama mengandung kata timnas, indonesia, aff, final, piala, kita, pemain, 2020, sudah. Topik kedua mengandung kata timnas, indonesia, aff, juara, pemain, bisa, leg, lawan, pspi. Topik ketiga mengandung kata

timnas, shin, final, piala, kita, juara, pemain, 2020, kasih. Topik keempat mengandung kata timnas, indonesia, final, piala, pemain, 2020, bisa, gak, masuk. Topik kelima mengandung kata timnas, indonesia, aff, final, piala, 2020, pelatih, menang, untuk.

Gambar 15 menunjukkan topik modelling positif yang sering dibicarakan oleh suporter tim nasional Indonesia di media sosial Twitter. Pada hasil visualisasi terdapat delapan topik yang membahas tim nasional Indonesia di media sosial Twitter. Setiap topik memiliki beberapa jumlah kata dan setiap kata memiliki nilai beta. Kata yang memiliki nilai beta diatas 0,03 dapat dijadikan patokan untuk membuat kalimat yang mengandung informasi mengenai topik positif apa saja yang paling banyak dibahas. Kalimat tersebut sebagai berikut.

1. Tim nasional Indonesia (**garuda**) mendapatkan **dukungan** suporter
2. Tim nasional Indonesia berisikan **pemain muda**
3. Tim nasional Indonesia **masuk** final piala AFF
4. **Shin** Tae Yong merupakan **pelatih** Tim nasional Indonesia
5. Tim nasional Indonesia mendapatkan **apresiasi** di **leg kedua** final AFF
6. Tim nasional Indonesia mendapatkan ucapan **terima kasih** atas **perjuangan** di piala AFF



Gambar 15. Visualisasi Topic Modelling Positif (Data yang telah diolah)

Gambar 4.10 menunjukkan topik modelling negatif yang sering dibicarakan oleh suporter tim nasional Indonesia di media sosial Twitter. Pada hasil visualisasi terdapat delapan topik yang membahas tim nasional Indonesia di media sosial Twitter. Setiap topik memiliki beberapa jumlah kata dan setiap kata memiliki nilai beta. Kata yang memiliki nilai beta diatas 0,03 dapat dijadikan patokan untuk membuat kalimat yang mengandung informasi mengenai topik negatif apa saja yang paling banyak dibahas. Kalimat tersebut sebagai berikut.

1. **Pemain** Tim nasional Indonesia yang bernama **Dedik**
2. Tim nasional Indonesia hanya diberi makan **nasi kotak**
3. Ketua Umum (**Ketum**) **PSSI Jangan masuk ruang ganti** pemain
4. Tim nasional Indonesia kalah saat bertanding melawan **Thailand**
5. Pemain Tim nasional Indonesia **Asnawi** melakukan provokasi terhadap Ramli



Gambar 16. Visualisasi *topic modelling* negatif (Data yang telah diolah)

D. Pembahasan

Data yang ada pada penelitian ini berhasil diolah, dikelompokkan, dianalisis, dan disimpulkan. Data diambil dari media sosial Twitter dalam rentang waktu 23 Desember 2021 – 1 Januari 2022. Data sebelum dilakukan preprocessing memiliki jumlah 10.000 data. Setelah dilakukan preprocessing dengan menggunakan Rapidminer, jumlah data menyusut menjadi 1.247 data. Preprocessing dilakukan dengan menggunakan operator yang tersedia di Rapidminer diantaranya read CSV, subprocess, filter examples, remove duplicate, transform cases, tokenize, filter tokens, filter stopwords, stem dan write CSV.

Penelitian ini menggunakan analisis sentimen terhadap data tweet yang ada. Hasil dari analisis sentimen menunjukkan hasil bahwa 62% dari total tweet memiliki sentimen positif dan 38% memiliki sentimen negatif. Sehingga dapat dikatakan bahwa persepsi suporter terhadap tim nasional Indonesia bisa dikatakan baik dikarenakan persentase dari sentimen positif lebih unggul jika dibandingkan dengan persentase dari sentimen negatif. Hasil dari pengujian analisis sentimen dengan memanfaatkan metode naive bayes dilakukan dengan menggunakan aplikasi Rapidminer. Hasil akurasi terhadap metode klasifikasi yang diterapkan sebesar 64,41%. Topic modelling untuk mengetahui topik apa saja yang sering dibicarakan oleh suporter tim nasional Indonesia di sosial media *Twitter* mengenai tim nasional Indonesia di ajang kejuaraan AFF. Kata kata positif yang sering muncul dalam topik yaitu Garuda, dukung, pemain, muda, masuk, Shin, pelatih, apresiasi, leg, kedua, terima, kasih dan perjuangan. Sedangkan kata negatif yang paling sering muncul dalam topik adalah pemain, Dedik, nasi, kotak, ketum, PSSI, jangan, masuk, ruang, ganti, Thailand dan Asnawi.

IV. Kesimpulan

Berdasarkan hasil penelitian analisis sentimen dan *topic modelling* dapat disimpulkan bahwa metode naive bayes dalam melakukan klasifikasi memiliki akurasi sebesar 64,41%. Sentimen yang terbentuk pada persepsi suporter tim nasional Indonesia di sosial media twitter didominasi oleh sentimen positif sebesar 62% dan 38% memiliki sentimen negatif. Topik yang dibahas mengenai tim nasional Indonesia pada sentimen positif secara keseluruhan merupakan dukungan suporter terhadap tim nasional Indonesia yang berisikan pemain muda dan dilatih oleh pelatih Shin Tae Yong serta apresiasi terhadap perjuangan karena tim nasional berhasil masuk babak final kejuaraan AFF Suzuki Cup 2020. Topik pada sentimen negatif banyak membahas mengenai kekalahan tim nasional di babak final, pemain yang hanya diberi nasi kotak, pemain tim nasional yang bernama dedik dan rencana ketua umum PSSI masuk ruang ganti pemain saat pertandingan final.

REFERENSI

- (2021). Persib Bandung, Persija Jakarta & Klub Indonesia Dengan Pengikut Media Sosial Terbanyak, [online] <https://www.goal.com/id/galeri>
- Agustina, L. S. S. (2020). Perilaku Berbasis Teknologi di Sosial Media pada Digital Nattive. *Journal An-Nafs: Kajian Penelitian Psikologi*, 5(1), 76-88. Retrieved from <https://ejournal.iai-tribakti.ac.id/>

- Alfanzar, A. I. (2019). Topic Modelling Skripsi menggunakan metode Latent Dirichlet Allocation (Doctoral dissertation, UIN Sunan Ampel Surabaya). Retrieved from <http://digilib.uinsby.ac.id/>
- Anggraini, R. A., Widagdo, G., Budi, A. S., & Qomaruddin, M. (2019). Penerapan Data Mining Classification untuk Data Blogger Menggunakan Metode Naïve Bayes. *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, 7(1), 47-51. Retrieved from <https://jurnal.untan.ac.id/>
- Asih, R. W. (2021). Sejarah Piala AFF: Bermula dari Piala Tiger hingga Sistem kandang-Tandang, [online]. <https://sport.bisnis.com>
- Azhari, M., Situmorang, Z., & Rosnelly, R. (2021). Perbandingan Akurasi, Recall, dan Presisi Klasifikasi pada Algoritma C4. 5, Random Forest, SVM dan Naive Bayes. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 640-651. Retrieved from <http://stmik-budidarma.ac.id/>
- Bernik, M. (2019). Analisis kualitas pelayanan untuk meningkatkan kepuasan konsumen: studi kasus gold's gym bandung. *Jurnal Manajemen Teori dan Terapan*, 12(3), 6-8. Retrieved from academia.edu
- Cahyono, A. S. (2016). Pengaruh media sosial terhadap perubahan sosial masyarakat di Indonesia. *Publiciana*, 9(1), 140-157. Retrieved from journal.unita.ac.id
- Chandani, V., Wahono, R. S., & Purwanto, P. (2015). Komparasi algoritma klasifikasi Machine Learning dan feature selection pada analisis sentimen review film. *Journal of Intelligent Systems*, 1(1), 56-60. Retrieved from <https://www.neliti.com/>
- David, W., & David, F. (2020). Analisis Sensori Lanjut Untuk Industri Pangan dengan R: Preference mapping dan survival analysis. Retrieved from <http://repository.bakrie.ac.id/>
- Deolika, A., Kusri, K., & Luthfi, E. T. (2019). Analisis Pembobotan Kata Pada Klasifikasi Text Mining. *JurTI (Jurnal Teknologi Informasi)*, 3(2), 179-184. Retrieved from <http://www.jurnal.una.ac.id/>
- Dihni, A. V. (2021). Inilah 10 Negara Dengan Pengguna Twitter Terbanyak Ada Indonesia, [online] <https://databoks.katadata.co.id>
- Fatoni, M. (2019). Daftar 10 kelompok suporter terbanyak di ASEAN versi Fox Sports Asia, [online] <https://jogja.tribunnews.com/>
- Federation Internationale de Football Association, [online] <https://www.fifa.com/fifa-world-ranking/men>
- Fitriyiah, S. N. J., Safriadi, N., & Pratama, E. E. (2019). Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes. *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, 5(3), 279-285. Retrieved from <https://jurnal.untan.ac.id/>
- Hozairi, H., Anwari, A., & Alim, S. (2021). Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes. *Network Engineering Research Operation*, 6(2), 133-144. Retrieved from <https://nero.trunojoyo.ac.id/>
- Hudaya, C. S., Fakhurroja, H., & Alamsyah, A. (2019). Analisis Persepsi Konsumen Terhadap Brand Go-Jek Pada Media Sosial Twitter Menggunakan Metode Sentiment Analysis dan Topic Modelling. *Jurnal Mitra Manajemen*, 3(6), 664-673. Retrieved from <http://e-jurnalmitramanajemen.com/>
- Indrawati. (2015). Metode Penelitian Manajemen dan Bisnis Konvergensi Teknologi Komunikasi dan Informasi. Bandung: Aditama
- Ipmawati, Joang., Kusri, dan Luthfi, E.T. (2017). Komparasi Teknik Klasifikasi Teks Mining pada Analisis Sentimen. *Indonesian Journal on Networking and Security*. 6(1): 28- 36.
- Iriawan, M. (2022). Nasib Shin Tae-Yong Dan Bongkar Mafia Bola, [online] <https://www.youtube.com/watch?v=dN8tcVkhGT4>
- Jailani, M. (2019). Hubungan status sosial ekonomi orang tua terhadap motivasi anak untuk berwirausaha. *Pedagogik: Jurnal Pendidikan*, 14(1), 35-42. Retrieved from journal.umpalangkaraya.ac.id
- Kamhar, M. Y., & Lestari, E. (2019). Pemanfaat Sosial Media Youtube Sebagai Media Pembelajaran Bahasa Indonesia DI Perguruan Tinggi. *Inteligensi: Jurnal Ilmu Pendidikan*, 1(2), 1-7. Retrieved from <https://jurnal.unitri.ac.id/>
- Kotler, P., & Keller, K. L. (2016). *Marketing Management 15th Global edition* (Global). Harlow: Pearson Education Limited.
- Mahardhika, Y. S., & Zuliarso, E. (2018). Analisis Sentimen Terhadap Pemerintahan Joko Widodo Pada Media Sosial Twitter Menggunakan Algoritma Naives Bayes Classifier. Retrieved from <https://unisbank.ac.id/>
- Manalu, E., Sianturi, F. A., & Manalu, M. R. (2017). Penerapan Algoritma Naive Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan Dan Jumlah Pemesanan Pada Cv. Papadan Mama Pastries. *Jurnal Mantik Penusa*, 1(2). Retrieved from <https://d1wqtxts1xzle7.cloudfront.net/>
- Mardianie, R. (2018). PENGARUH PROFESIONALISME KERJA, PENGHARGAAN DAN PENGAWASAN KERJA TERHADAP KINERJA TENAGA KONTRAK KEBERSIHAN BIDANG PERTAMANAN DAN KEBERSIHAN DI DINAS PEKERJAAN UMUM, PENATAAN RUANG, PERUMAHAN DAN KAWASAN PERMUKIMAN KABUPATEN KAPUAS. *Jurnal Ilmiah Ekonomi Bisnis*, 4(2). Retrieved from <https://ejournal.stiepancasetia.ac.id/>
- Masrury, R. A., & Alamsyah, A. (2019, July). Analyzing tourism mobile applications perceived quality using sentiment analysis and topic modeling. In 2019 7th international conference on information and communication technology (ICoICT) (pp. 1-6). IEEE

- Mukti, T. B. (2021). Piala AFF 2020 Dipastikan Tak Ditunda Tahun Ini Tapi Ada Perubahan Format, [online]. <https://timnas.skor.id/>
- Naury, C., Fudholi, D. H., & Hidayatullah, A. F. (2021). Topic modelling pada sentimen terhadap headline berita online berbahasa indonesia menggunakan LDA dan LSTM. *Jurnal Media Informatika Budidarma*, 5(1), 24-33. Retrieved from <https://www.ejournal.stmik-budidarma.ac.id/>
- Novianto, H. (2020). Dari Mana Klub Sepak Bola Meraih Penghasilan, [online] <https://boladunia.skor.id/>
- Nugroho, Y. S. (2014). Data Mining Menggunakan Algoritma Naive Bayes Untuk Klasifikasi Kelulusan Mahasiswa Universitas Dian Nuswantoro. Dian Nuswantoro Fakultas Ilmu Komputer Skripsi. Retrieved from <http://academia.edu>
- Panjaitan, W., Lie, D., Butarbutar, M., & Inrawan, A. (2018). Pengaruh Kepuasan Pelanggan dan Persepsi Pelanggan terhadap Loyalitas Pelanggan pada Ud Pn Silitonga Serbelawan. *Maker: Jurnal Manajemen*, 4(1), 83-91. Retrieved from <https://maker.ac.id/>
- Persatuan Sepak Bola Seluruh Indonesia, [online] <https://www.pssi.org/>
- Pradigdo, B. G. (2021). Perubahan Format Dalam Perjalanan Piala AFF Mulai Home Tournament 2 Tuan Rumah Dan Home And Away, [online]. <https://www.bola.com/>
- Priyanto, A., & Ma'arif, M. R. (2018). Implementasi Web Scrapping dan Text Mining untuk Akuisisi dan Kategorisasi Informasi dari Internet (Studi Kasus: Tutorial Hidroponik). *Indonesian Journal of Information Systems*, 1(1), 25-33. Retrieved from <https://ojs.uajy.ac.id/>
- Pulakiang, A. R., Wiranatha, I. M. N., & Batmetan, J. R. (2019). ANALISIS KUALITAS LAYANAN E-COMMERCE MENGGUNAKAN TWITTER API (Studi Kasus: Tokopedia, Lazada dan Bukalapak). *IJIS-Indonesian Journal On Information System*, 4(1). Retrieved from <http://www.ijiswiratama.org/>
- Puspita, R., & Widodo, A. (2021). Perbandingan Metode KNN, Decision Tree, dan Naive Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS. *Jurnal Informatika Universitas Pamulang*, 5(4), 646-654. Retrieved from scholar.archive.org
- Rafiah, K. K. (2019). Analisis pengaruh kepuasan pelanggan dan kepercayaan pelanggan terhadap loyalitas pelanggan dalam berbelanja melalui E-commerce di Indonesia. *Al Tijarah*, 5(1), 46-56. Retrieved from <https://ejournal.unida.gontor.ac.id/>
- Ratnawati, F. (2018). Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter. *INOVTEK Polbeng-Seri Informatika*, 3(1), 50-59. Retrieved from <http://ejournal.polbeng.ac.id/>
- Rec.Sport.Soccer Statistics Foundation, [online] <http://www.rsssf.com/>
- Ridwan, M. (2022). DAFTAR JUARA Piala AFF 1996-2020, [online]. <https://www.goal.com/>
- Rijal, M., Jumadi, J., & Tenri Awaru, A. O. (2021). Solidaritas Fans Klub Kota Makassar (Studi: Milanisti Sezione Makassar). *Phinisi Integration Review*, 4(3), 414-422. Retrieved from <http://eprints.unm.ac.id/>
- Rozi, I. F., Hamdana, E. N., & Alfahmi, M. B. I. (2018). Pengembangan Aplikasi Analisis Sentimen Twitter Menggunakan Metode Naive Bayes Classifier (Studi Kasus SAMSAT Kota Malang). *Jurnal Informatika Polinema*, 4(2), 149-149. Retrieved from <http://jip.polinema.ac.id/>
- Rudi, A. (2018). Jumlah penonton Indonesia vs Jepang jadi rekor baru di Piala Asia U-19, [online] <https://bola.kompas.com/>
- Rustiana, D., & Rahayu, N. (2017). Analisis Sentimen Pasar Otomotif Mobil: Tweet Twitter Menggunakan Naive Bayes. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 8(1), 113-120. Retrieved from <http://jurnal.umk.ac.id/>
- Saefuloh, D. (2020). Rekognisi terhadap peluang bisnis online melalui media sosial dan hubungannya dengan minat berwirausaha: studi kasus mahasiswa Politeknik Negeri Bandung. *Jurnal Bisnis Dan Kewirausahaan*, 16(1), 12-23. Retrieved from <https://ojs.pnb.ac.id/>
- Salsabila, K. D. A., & Trianasari, N. (2021). Analisis Persepsi Produk Kosmetik Menggunakan Metode Sentiment Analysis dan Topic Modeling (Studi Kasus: Laneige Water Sleeping Mask). *Jurnal Teknologi dan Manajemen Informatika*, 7(1), 1-9.
- Sasmita, D., Ariyanti, M., & Febrianta, M. Y. (2021). Analisis Kualitas Layanan Pada Platform E-commerce Indonesia Menggunakan Topic Modeling Dan Analisis Sentimen (studi Kasus: Tokopedia, Shopee, Bukalapak). *eProceedings of Management*, 8(1). Retrieved from <https://openlibrarypublications.telkomuniversity.ac.id/>
- Semuel, H., & Setiawan, K. Y. (2018). Promosi melalui sosial media, brand awareness, purchase intention pada produk sepatu olahraga. *Jurnal Manajemen Pemasaran*, 12(1), 47-52. Retrieved from http://203.189.121.7/~puslit2_ejournal/ejournal/index.php
- Siregar, N. (2018). Analisis faktor-faktor yang mempengaruhi minat beli konsumen dalam menggunakan indihome sebagai penyedia jasa internet di kota medan (studi kasus kantor plaza telkomcabang iskandar muda no. 35 medan baru). *Jumant*, 7(1), 65-76. Retrieved from <https://journal.pancabudi.ac.id/>
- Sudiantoro, A. V., & Zuliarso, E. (2018). Analisis sentimen twitter menggunakan text mining dengan algoritma Naive Bayes Classifier. Retrieved from <https://www.unisbank.ac.id/>

- Sumarno, S., & Irianto, D. P. (2019). Studi eksplorasi klub sepakbola pada Liga 2 Indonesia 2018: Bagaimana cara bertahan dalam keterbatasan?. *Jurnal Keolahragaan*, 7(2), 104-115. Retrieved from <https://journal.uny.ac.id/>
- Tampubolon, M. R., Yunus, M., & Liana, L. (2019). Pengaruh Persepsi Konsumen terhadap Keputusan Pembelian pada Perum Bulog sub Divre Pematangsiantar. *Jesya (Jurnal Ekonomi dan Ekonomi Syariah)*, 2(2), 462-467. Retrieved from stiealwashliyahsibolga.ac.id
- Virantika, D. (2021). 12 Suporter Paling Fanatik dan Berpengaruh di Sepakbola Indonesia, Nomor 1 Fans di Luar Negeri, [online] <https://bola.okezone.com>
- Waskito, L. A., Lhaksana, K. M., & Murdiansyah, D. T. (2019). Analisis Sentimen Terhadap Pemilihan Presiden Indonesia 2019 Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes. *eProceedings of Engineering*, 6(2). Retrieved from <https://openlibrarypublications.telkomuniversity.ac.id/>
- Wati, R. (2020). Penerapan Algoritma Naive Bayes Dan Particle Swarm Optimization Untuk Klasifikasi Berita Hoax Pada Media Sosial. *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, 5(2), 159-164. Retrieved from <http://ejournal.nusamandiri.ac.id/>