ANALISIS *SENTIMENT* MASYARAKAT TERHADAP *SEA GAMES ESPORTS* PADA *TWITTER* DENGAN MENGGUNAKAN ALGORITMA *ADABOOST*

SKRIPSI

OLEH:

MUHAMMAD SYIFA

188160025

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MEDAN AREA
MEDAN
2024

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang
---
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24

Access From (repository.uma.ac.id)13/9/24

ANALISIS *SENTIMENT* MASYARAKAT TERHADAP *SEA GAMES ESPORTS* PADA *TWITTER* DENGAN MENGGUNAKAN ALGORITMA *ADABOOST*

PROPOSAL SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana di Fakultas Teknik

Universitas Medan Area

OLEH :

MUHAMMAD SYIFA 188160025

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MEDAN AREA

MEDAN

2024

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Document Accepted 13/9/24

HALAMAN PENGESAHAN

Judul Skripsi : Analisis *Sentiment* Masyarakat Terhadap *Sea Games eSports* Pada *Twitter* Dengan Menggunakan Algoritma *Adaboost*

Nama : Muhammad Syifa

NPM : 188160025

Fakultas : Teknik

Disetujui Oleh :

Komisi Pembimbing

Nurul Khairina, S.Kom, M.Kom

Dr. Dan Suriatno, ST., MT
Dekan Fakultas Teknik

Rizki Muliono, S. Kom., M. Kom
Ka. Prodi Teknik Informatika

Tanggal lulus: 03 April 2024

© Hak Cipta Di Lindungi Undang-Undang

Document Accepted 13/9/24

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Access From (repository.uma.ac.id)13/9/24

HALAMAN PERNYATAAN

Saya menyatakan bahwa skripsi yang saya susun, sebagai syarat memperoleh gelar serjana merupakan hasil karya tulis saya sendiri. Adapun bagian-bagian tertentu dalam penulisan skripsi ini yang saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas sesuai dengan norma, kaidah, dan etika penulisan ilmiah.

Saya bersedia menerima sanksi pencabutan gelar akademik yang saya peroleh dan sanksi-sanksi lainnya dengan peraturan yang berlaku, apabila di kemudian hari ditemukan adanya plagiat dalam skripsi ini.

Medan, 30 Agustus 2024

Munammad Syifa
NPM 188160025

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24

Access From (repository.uma.ac.id)13/9/24

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR/SKRIPSI/TESIS UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Medan Area, saya yang bertanda tangan di bawah ini:

Nama	: Muhammad Syifa
NPM	: 188160025
Program Studi	: Teknik Informatika
Fakultas	: Teknik
Jenis Karya	: Skripsi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Medan Area **Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty- Free Right)** atas karya ilmiah saya yang berjudul :

Analisis *Sentiment* Masyarakat Terhadap *Sea Games eSports* Pada *Twitter* Dengan Menggunakan Algoritma *Adaboost*

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Medan Area berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan tugas akhir/skripsi/tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Medan

Pada tanggal 20 Februari 2024

Yang menyatakan

(Muhammad Syifa)

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24

Access From (repository.uma.ac.id)13/9/24

ABSTRAK

Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap Sea Games eSports yang tercermin dalam percakapan di platform Twitter. Dua permasalahan utama yang menjadi fokus adalah bagaimana sentimen masyarakat terhadap acara Sea Games eSports dan apakah tanggapan mereka cenderung positif atau negatif terhadap acara tersebut. Metode penelitian yang digunakan adalah Algoritma Adaboost, sebuah teknik pada data mining yang bertujuan meningkatkan akurasi klasifikasi. Adaboost digunakan untuk menganalisis sentimen dari data percakapan di Twitter dengan menggunakan feature selection untuk memilih fungsi klasifikasi yang lemah, kemudian menggabungkannya menjadi fungsi klasifikasi yang baru. Hasil penelitian menunjukkan evaluasi tertinggi tercapai pada pengujian ke-2 dengan penggunaan data training sebesar 90% dan testing sebesar 10%, yang menghasilkan akurasi sebesar 98%. Analisis sentimen terhadap Sea Games eSports menunjukkan bahwa mayoritas pengguna twitter mengekspresikan sentimen positif sebanyak 111 (95,7%), sementara sentimen negatif hanya sebanyak 5 (4,3%). Hal ini mengindikasikan bahwa masyarakat cenderung memberikan tanggapan yang positif terhadap acara Sea Games eSports. Penelitian ini memberikan gambaran bahwa analisis sentimen dengan menggunakan Algoritma Adaboost mampu memberikan hasil akurat dalam memetakan tanggapan masyarakat terhadap acara Sea Games eSports di Twitter. Implikasinya, adanya dukungan yang kuat dari masyarakat terhadap acara ini dapat menjadi dasar penting dalam pengembangan dan promosi lebih lanjut dari Sea Games eSports.

Kata Kunci : Sentimen, Esport, Adaboost, Klasifikasi

Document Accepted 13/9/24

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Access From (repository.uma.ac.id)13/9/24

ABSTRACT

This research aimed to analyse the public sentiment towards the Sea Games eSports as reflected in conversations on the Twitter platform. The two main issues in focus were how the public sentiment towards the Sea Games eSports event and whether their responses tended to be positive or negative towards the event. The research method used was the Adaboost Algorithm, a technique in data mining that aimed to improve classification accuracy. Adaboost was used to analyse sentiment from Twitter conversation data by using feature selection to select weak classification functions, then combining them into a new classification function. The results showed that the highest evaluation was achieved in the 2nd test with the use of 90% training data and 10% testing data, which resulted in an accuracy of 98%. Sentiment analysis of the Sea Games eSports showed that the majority of twitter users expressed positive sentiments as many as 111 (95.7%), while negative sentiments were only 5 (4.3%). This indicated that people tended to give positive responses to the Sea Games eSports event. This research illustrated that sentiment analysis using the Adaboost Algorithm was able to provide accurate results in mapping public responses to the Sea Games eSports event on Twitter. The implication is that the strong support from the public towards this event can be an important basis for further development and promotion of Sea Games eSports.

Keywords: *Sentiment, Esport, Adaboost, Classification*

24/05/2024

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang
---
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24

Access From (repository.uma.ac.id)13/9/24

RIWAYAT HIDUP

Penulis dilahirkan di Bukit Lawang pada tanggal 24 Juli 2000 sebagai anak ke-1 (satu) dari 1 (satu) bersaudara dari pasangan yang Bernama Saefudin dan Atmini. Saat ini penulis tinggal di Perkebunan Bukit Lawang, Kec. Bahorok, Kab. Langkat, Sumatera Utara

PadaTahun 2018 Penulis lulus dari SMA Negeri 1 Bahorok, lalu pada tahun 2018 Penulis terdaftar sebagai mahasiswa Fakultas Teknik Prodi Teknik Informatika Universitas Medan Area

Pada tahun 2021 penulis melaksanakan Kerja Praktek (KP) di Lembaga konservasi lingkungan hidup Nirlaba JUNAS (Jungle Nature Survival) Kec. Bahorok, Kab. Langkat, Sumatera Utara

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

KATA PENGANTAR

Puji syukur kita panjatkan kehadirat Allah SWT yang telah memberikan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan judul "Analisis Sentimen Masyarakat terhadap *Sea Games eSports* pada *Twitter* dengan menggunakan Algoritma *Adaboost*." Penulisan skripsi ini sebagai salah satu syarat untuk menyelesaikan pendidikan di tingkat Sarjana Komputer (S.Kom) Jurusan Teknik Informatika, Fakultas Teknik, Universitas Medan Area

Skripsi ini merupakan hasil kerja keras, dedikasi, dan dukungan dari berbagai pihak yang tidak dapat penulis sebutkan satu per satu. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Dadan Ramdan, M.Eng., M.Sc. Berkat semangat dan masukannya saat saya menjalankan PKKMB yang memotivasi saya sehingga dapat menyelesaikan tugas akhir atau penelitian saya saat ini.
2. Bapak Dr. Eng. Supriatno, ST, MT selaku Dekan Fakultas Teknik Universitas Medan Area.
3. Bapak Rizki Muliono, S.Kom, M.Kom selaku Kaprodi yang telah membimbing saya dalam menentukan judul penelitian serta memotivasi saya agar terus mengerjakan skripsi saya hingga selesai.
4. Ibu Nurul Khairina, S.Kom, M.Kom, selaku pembimbing skripsi, atas bimbingan, arahan, dan waktunya yang telah diberikan dalam mengarahkan penelitian ini sehingga menjadi lebih baik.
5. Teman-teman sejawat yang telah berkontribusi dalam berbagai cara, mulai dari diskusi ilmiah hingga dukungan moral selama perjalanan penelitian ini.
6. Orangtua dan keluarga yang selalu memberikan doa, semangat, dan dukungan moral dalam setiap langkah penulis selama proses penelitian dan penulisan skripsi ini.
7. Semua responden yang telah berpartisipasi dalam penelitian ini dengan memberikan data melalui Twitter. Kontribusi mereka sangat berarti dalam mengevaluasi sentimen masyarakat terhadap Sea Games eSports.
8. Seluruh sivitas akademika Universitas Medan Area yang telah menciptakan lingkungan akademik yang mendukung penelitian dan pengembangan ilmu pengetahuan.
9. Seluruh jajaran pegawai/staff Teknik Universitas Medan Area yang telah melaksanakan tugasnya serta mengurus dokumen-dokumen mahasiswa Fakultas Teknik dengan baik.

Penulis menyadari bahwa penelitian ini masih jauh dari sempurna. Oleh karena itu, masukan dan saran yang membangun sangat diharapkan untuk perbaikan di masa mendatang. Skripsi ini diharapkan dapat memberikan kontribusi kecil terhadap pengembangan ilmu pengetahuan, khususnya dalam bidang analisis sentimen dan penggunaan algoritma Adaboost dalam konteks eSports.

Document Accepted 13/9/24
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Akhir kata, semoga skripsi ini dapat bermanfaat bagi pembaca yang berminat untuk melanjutkan penelitian di bidang yang sama. Semoga penelitian ini juga dapat menjadi salah satu referensi dalam pengembangan analisis sentimen di media sosial.

Medan, 30 Agustus 2024

Muhammad Syifa
NPM 188160025

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

DAFTAR ISI

ABSTRAK ..... **v**
ABSTRACT ..... **vi**
RIWAYAT HIDUP ..... **vii**
KATA PENGANTAR ..... **viii**
DAFTAR GAMBAR ..... **xii**
DAFTAR TABEL ..... **xiii**
BAB I PENDAHULUAN **1**
1.1. Latar Belakang 1
1.2. Perumusan Masalah 4
1.3. Tujuan Penelitian 4
1.4. Batasan Penelitian 5
1.5. Manfaat Penelitian 5
BAB II TINJAUAN PUSTAKA **7**
2.1. Analisis Sentimen 7
2.2. Twitter 7
2.3. Sea Games 8
2.4. Olahraga (Sports) 8
2.4.1. Olahraga Elektronik (*eSports*) 9
2.5. Text Mining 10
2.6. Algoritma *Adaboost* 10
2.7. Python 11
2.8. Confusion Matrix 12
2.9. *Term Frequncy Inverse Document Frequency* (TF-IDF) 14
2.10. Penelitian Terdahulu 17
BAB III METODE PENELTIAN **20**
3.1. Waktu dan Tempat Penelitian 20
3.2. Metodologi Penelitian 20
3.2.1. Pengumpulan data (*Scrapping Data*) 21
3.2.2. *Text Preproccessing* 23
a) *Case Folding* 24
b) *Tokenization* 24
c) *Stopwords* 26

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

d) *Stemming* 27
3.2.3. *Sentiment Intensity Analyzer* 28
3.2.4. *TF-IDF* 29
3.3. Algoritma *Adaboost* 31
BAB IV HASIL DAN PEMBAHASAN 35
4.1. Hasil 35
4.1.1. Pengumpulan data 35
4.1.2. Text Preprocessing 36
4.1.4. Feature Extraction (TF-IDF) 43
4.1.5. Algoritma Adaboost 44
4.2. Pembahasan 54
4.2.1. Analisis Hasil 54
4.2.2. Interpretasi Temuan 54
4.3.3. Perbandingan dengan Penelitian terdahulu 55
4.3.4. Keterbatasan Penelitian 56
BAB V KESIMPULAN DAN SARAN 57
5.1 Kesimpulan 57
5.2 Saran 57
DAFTAR PUSTAKA 58

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

DAFTAR GAMBAR

Gambar 3. 1. Alur Proses Penelitian 20
Gambar 3. 2. Portal developer twitter 21
Gambar 3. 3. Scrape Data Menggunakan API 22
Gambar 3. 4. Alur algoritma adaboost 31
Gambar 4. 1. Tampilan Data 35
Gambar 4. 2. Tahapan Case Folding 36
Gambar 4. 3. Tahapan Tokenization 37
Gambar 4. 4. Tahapan Stopwords 39
Gambar 4. 5. Tahapan Stemming 40
Gambar 4. 6. Source Code Sentiment Intensity Analyzer 42
Gambar 4. 7. Hasil Sentiment Intensity Analyzer 42
Gambar 4. 8. Sentimen aktual dari Sentiment Intensity Analyzer 43
Gambar 4. 9 Feature Exctracion TF-IDF 44
Gambar 4. 10 Algoritma Adaboost 44
Gambar 4. 11 Confusion Matrix data testing 40% 47
Gambar 4. 12 Confusion Matrix data testing 30% 48
Gambar 4. 13 Confusion Matrix data testing 20% 49
Gambar 4. 14 Confusion Matrix data testing 10% 50
Gambar 4. 15 Visualisasi Sentimen dengan Data Testing 40% 51
Gambar 4. 16 Visualisasi Sentimen dengan Data Testing 30% 52
Gambar 4. 17 Visualisasi Sentimen dengan Data Testing 20% 53
Gambar 4. 18 Visualisasi Sentimen dengan Data Testing 10% 53

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix 13
Tabel 2. 2 Penelitian Terdahulu 17
Tabel 3. 1 Data Sampel 23
Tabel 3. 2 Text to Case Folding 24
Tabel 3. 3 Case Folding to Tokenization 25
Tabel 3. 4 Tokenization to Stopwords 26
Tabel 3. 5 Stopwords to Steeming 27
Tabel 3. 6 Sentiment Intensity Aanlyzer 28
Tabel 3. 7 TF-IDF 29
Tabel 3. 8. Data sampel perhitungan manual 32
Tabel 3. 9. Data sampel setelah dilatih menggunakan algoritma adaboost 32
Tabel 3. 10. Output algoritma adaboost 33
Tabel 4. 1 Hasil Case Folding dengan 10 data 36
Tabel 4. 2 Hasil Tokenization dengan 10 data 37
Tabel 4. 3 Hasil Stopwords dengan 10 data 39
Tabel 4. 4 Hasil Stemming dengan 10 data 40
Tabel 4. 5 Hasil 10 data dengan algoritma adaboost 44
Tabel 4. 6 Evaluasi dengan Confusion Matrix 46

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

BAB I
PENDAHULUAN

1.1. Latar Belakang

Analisis sentimen adalah cara untuk mengetahui apakah suatu teks mengandung perasaan positif atau negatif. Ada beberapa cara teknis untuk melakukannya, seperti menggunakan mesin pembelajaran, kosakata, aturan, atau model statistik. (Septian dkk., 2019). Analisis sentimen berguna dalam mengidentifikasi perasaan dalam teks, seperti ketika perusahaan ingin mengetahui bagaimana pelanggan merespons produk atau layanan mereka. Dengan menggunakan teknologi ini, perusahaan dapat secara otomatis menganalisis umpan balik pelanggan dari survei atau media sosial, dan memahami apa yang membuat pelanggan senang atau kecewa. Contohnya, dengan menganalisis lebih dari 4.000 ulasan produk secara otomatis, perusahaan dapat mengetahui apakah pelanggan puas dengan harga dan layanan pelanggan mereka. Selain itu, dengan memantau sentimen media secara real-time, perusahaan dapat segera mengidentifikasi pelanggan yang tidak puas dan merespons dengan cepat. (Wibowo Patrick Trusto Jati, 2021)

Olahraga elektronik (esports) merupakan bentuk kompetisi video game yang telah menjadi semakin populer dan profesional selama beberapa dekade terakhir. Industri esports global diprediksi akan menghasilkan pendapatan 1,084 miliar USD atau sekitar 16,9 triliun rupiah. (Lestari dkk., 2023).

E-Sports atau yang biasa disebut olahraga elektronik masih banyak yang membicarakan antara olahraga atau bukan. Banyaknya pro dan kontra yang membahas E-Sports ini, namun terlepas dari semua itu E-Sports semakin

Document Accepted 13/9/24
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

berkembang di Indonesia. Semakin berkembangnya E-Sports atau olahraga elektronik telah menjadikan olahraga ini masuk ke dalam cabang olahraga baru yang dipertandingkan di event Sea Games 2019 yang diselenggarakan di kota Jakarta dan Palembang. Meskipun E-Sports masih dalam perkembangannya, namun hal tersebut cukup membuktikan bahwasanya E-Sports merupakan cabang olahraga baru yang eksis dan cukup banyak diakui, sehingga olahraga ini mempunyai pengikut cukup banyak di Indonesia maupun di dunia. (Renjani dkk., 2023). E-sport merupakan aktivitas yang unik, dimana semua memerlukan hubungan yang harmonis dan ideal antara proses berfikir, emosi dan gerakan kelincahan tangan. Apalagi dengan berkembangnya teknologi sekarang olahraga juga tidak mau ketinggalan dalam kemajuan teknologi saaat ini, maka dengan adanya penelitian ini peneliti bisa mengatahui apa yang ada di masyarakat saat ini, fenomena olahraga elektronik ini sejauh mana di lingkungan masyarakat. (Khudzaifah dkk., 2023)

Twitter adalah layanan jejaring sosial (media umum) dan microblogging yang memungkinkan penggunanya mengirim serta membaca pesan hingga 280 karakter, yang disebut menciak.. Twitter mencapai popularitasnya ketika diadakan Konferensi South oleh Southwest Interactive (SXSWi) tahun 2007. Selama acara berlangsung, jumlah pengguna Twitter tumbuh dari 20.000 kicauan (Tweet) menjadi 60.000 kicauan (Tweet) per hari. Menariknya, pesan Twitter pertama dikirim dari luar Bumi oleh astronot NASA TJ Creamer pada 22 Januari 2010. Twitter adalah salah satu media umum yang cukup di minati warga indonesia. Menurut data tahun 2019, lebih 22,8% pengguna internet di Indonesia menggunakanTwitter. Jika ada sesuatu yang sedang berbicara (tren), Twitter bisa

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

menjadi acuan. Semua topik serius dan remeh di bahas di Twitter, mulai asal politik, kepercayaan, hiburan, skandal hingga kisah-kisah menarik dan inspiratif.. (Arifin Dian Rudi, 2023)

Dengan adanya platform media sosial yang mampu mewadahi ribuan opini publik, dapat dimanfaatkan untuk mencari sebuah informasi terkait hal-hal tertentu, dengan menganalisis opini-opini publik dengan analisis sentimen. Analisis sentimen berguna untuk menentukan komentar atau opini memiliki kecenderungan sentimen negatif atau positif yang dapat dijadikan sebagai acuan dalam peningkatan suatu layanan, ataupun kualitas produk. (Febriyani & Februariyanti, 2022)**.** Mendeteksi suasana hati yang buruk adalah tujuan yang dikejar oleh banyak peneliti. Alat analisis sentimen yang disebutkan menawarkan cara untuk menentukan suasana hati berdasarkan komunikasi berbasis teks. Selain itu, alat analisis sentimen juga telah diterapkan dalam skenario aplikasi lainnya, termasuk pengembangan saran perbaikan untuk kode atau rekomendasi untuk paket dan perpustakaan *Software Engineering* yang lebih baik. (Sejati dkk., 2023)

Adaboost adalah merupakan salah satu algoritma yang dapat dipergunakan dalam melakukan identifikasi dan klasifikasi (Wahyu dkk., 2023). Adaboost dapat melakukan optimasi boosting dan dapat digabungkan dengan algoritma lainnya, seperti algoritma *Naïve Bayes* (Rizal Ramli & Sulastri, 2023). Adaboost adalah singkatan dari Adaptive Boosting, algoritma machine learning yang dikembangkan oleh Yoav Freund dan Robert Schapire. Adaboost secara teoritis dapat digunakan secara signifikan untuk mengurangi kesalahan dari beberapa algoritma pembelajaran yang secara konsisten meningkatkan kinerja classifier. Algoritma Adaboost dirancang khusus untuk memecahkan masalah klasifikasi dan membantu

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

meningkatkan akurasi algoritma yang lemah. Adaboost merupakan ensemble learning yang biasa digunakan untuk meningkatkan algoritma. Boosting dapat dikombinasikan dengan algoritma klasifikasi lain untuk meningkatkan kinerja klasifikasi. Tentu saja, menggabungkan beberapa model secara intuitif berguna ketika model berbeda satu sama lain (Novaneliza dkk., 2023). Hasil pengujian dan implementasi metode AdaBoost untuk mengklasifikasikan tweet menunjukkan bahwa penelitian ini berhasil menggunakan algoritma ini untuk mengklasifikasikan tweet. (Wahyu dkk., 2023)

Permasalahan kenapa dibuatnya analisis sentiment adalah untuk memudahkan dalam mengetahui sentimen dari setiap opini dari maupun pendapat pengguna *Twitter* mengenai tren yang sedang terjadi .Pada setiap kalimat perlu dibuatkan suatu sistem untuk menganalisis sentimen. (Septian dkk., 2019)

1.2. Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, maka dapat dirumuskan masalah yang ingin diselesaikan dalam penelitian ini, yaitu:

1. Bagaimana sentimen masyarakat terhadap *Seagames eSports* yang terungkap di *Twitter*?
2. Apakah tanggapan masyarakat cenderung positif atau negatif terhadap acara tersebut?

1.3. Tujuan Penelitian

Tujuan dari Penelitian ini adalah :

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

1. Membuat pemetaan tanggapan publik terhadap acara tersebut berdasarkan analisis sentimen.
2. Menguji keakuratan Algoritma *Adaboost* dalam melakukan klasifikasi sentimen.

1.4. Batasan Penelitian

Penelitian ini memiliki beberapa batasan, yaitu:

1. Data yang digunakan terbatas pada cuitan di *Twitter yang terkait dengan Seagames eSports.*
2. Penelitian ini fokus pada analisis sentimen masyarakat, dan tidak mencakup aspek lain seperti analisis tren atau demografi pengguna *Twitter*.
3. Algoritma *Adaboost* digunakan sebagai metode klasifikasi sentimen, namun tidak mencakup perbandingan dengan metode klasifikasi lainnya.
4. Jumlah data yang digunakan sebanyak 1153 data pengguna *twitter* dengan cara *Scrape* data dengan querry '*Seagames eSports* 2023' dengan menggunakan API *(Application Programming Interface) Twitter* yang mana berasal dari *website developer.twiiter.com*
5. *Python* merupakan bahasa pemrograman yang digunakan pada penelitian ini

1.5. Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Sebagai masukan bagi penyelenggara acara dalam meningkatkan pengalaman peserta dan penonton berdasarkan tanggapan masyarakat.

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Document Accepted 13/9/24

2. Sebagai acuan bagi sponsor dalam mengevaluasi potensi eksposur merek dan keterlibatan dengan acara *eSports*.
3. Memberikan pemahaman lebih dalam tentang persepsi dan preferensi masyarakat terhadap *Seagames eSports* di *Twitter*.

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Document Accepted 13/9/24

BAB II

TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen merupakan proses evaluasi teks yang menggunakan teknik analisis teks untuk mengidentifikasi emosi dari penulisnya(Ahmed dkk., 2021)

Komputasi atas opini-opini, sentimen, dan ekspresi emosional dalam teks juga dikenal sebagai analisis sentiment (Amrullah dkk., 2020)

Teknik ini bertujuan mengekstrak data teks guna memperoleh informasi mengenai sentimen yang dapat bersifat positif, netral, atau negatif. Pengguna internet seringkali menggunakan analisis sentimen di media sosial untuk memberikan penilaian atau ungkapan opini pribadi. (Sari & Wibowo, 2019)

2.2. Twitter

Twitter adalah suatu wadah media sosial yang memfasilitasi pengguna dalam berbagi informasi, pemikiran, dan pandangan dengan cepat dan mudah, serta menjalin interaksi dengan masyarakat secara luas. (Juliyanto & Rusdi, 2019)

Media sosial *Twitter* merupakan salah satu platform penghubung yang diminati oleh masyarakat global. Fenomena ini dapat terlihat dari pertumbuhan jumlah pengguna *Twitter* yang signifikan di seluruh dunia, termasuk di Indonesia. Pada tahun 2016, Twitter mencatat memiliki 313 juta pengguna aktif setiap bulan. Pengguna *Twitter* aktif terlibat dalam menyampaikan berita terkini atau memberikan komentar terkait topik utama yang sedang mendominasi perbincangan global. Topik-topik tersebut yang mendapatkan banyak perhatian dan komentar dari

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Document Accepted 13/9/24

pengguna *Twitter* sering kali dapat memicu kontroversi atau menjadi tren topik di berbagai media sosial, terutama *Twitter.* (Darwis dkk., 2021)

2.3. Sea Games

Pesta Olahraga Asia Tenggara, atau yang biasa disebut *SEA Games*, merupakan kompetisi multi-olahraga yang diselenggarakan setiap dua tahun dan melibatkan 11 negara di kawasan Asia Tenggara. Peraturan pertandingan *SEA Games* diatur oleh Federasi Pesta Olahraga Asia Tenggara (*Southeast Asian Games Federation*), dengan pengawasan dari Komite Olimpiade Internasional (IOC) dan Dewan Olimpiade Asia (OCA). (Hidayatulloh & Susanto, 2022)

2.4. Olahraga (Sports)

Olahraga memiliki keterkaitan yang erat dengan kehidupan manusia, di mana pun dan kapan pun mereka berada. Hal ini karena olahraga dianggap sebagai salah satu kebutuhan yang esensial bagi setiap individu. Peran olahraga sangat signifikan dalam kehidupan manusia, dan dalam era modern saat ini, aktivitas olahraga tidak hanya dianggap sebagai pekerjaan khusus, tetapi juga sebagai hiburan, kegiatan rekreasi, sumber mata pencaharian, pemeliharaan kesehatan, dan bagian dari warisan budaya. Pentingnya olahraga juga diakui dalam Undang-Undang Republik Indonesia Nomor 3 tahun 2005, khususnya dalam Bab II pasal IV, yang menetapkan sistem keolahragaan nasional. Menurut undang-undang tersebut, tujuan utama dari keolahragaan nasional adalah untuk menjaga dan meningkatkan kesehatan, kebugaran, prestasi, kualitas manusia, serta menanamkan nilai moral, sportivitas,

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Document Accepted 13/9/24

disiplin, mempererat persatuan dan kesatuan bangsa, memperkokoh ketahanan nasional, dan mengangkat martabat bangsa. (Adek Saputra, 2019)

2.4.1. Olahraga Elektronik (*eSports*)

Esport merupakan singkatan dari Olahraga Elektronik, dimana istilah "elektronik" merujuk pada sesuatu yang beroperasi menggunakan sejumlah komponen kecil, terutama mikrochip dan transistor, yang mengendalikan arus listrik. Sedangkan "olahraga" merujuk pada kegiatan fisik dan keterampilan yang dilakukan oleh individu atau kelompok dalam persaingan, baik antar individu maupun antar kelompok, hingga sebagai bentuk hiburan.(Nugraha Hidayat & Soenyoto, 2020)

eSports adalah singkatan dari olahraga elektronik atau "electronic sport," di mana permainan video digunakan sebagai arena pertandingan. Kompetisi ini diadakan secara daring dengan skema multipemain yang memungkinkan partisipasi dari individu di seluruh dunia. Karena kemudahan aksesnya, *eSports* telah menjadi cabang olahraga yang sangat diminati oleh berbagai kelompok usia. Tingginya minat masyarakat terhadap *eSports* mendorong perusahaan-perusahaan game untuk menyelenggarakan kejuaraan, baik tingkat regional maupun internasional, dengan hadiah yang menggiurkan. Dampaknya, tim-tim profesional muncul dan menggeluti *eSports* sebagai pekerjaan dengan imbalan finansial yang tinggi. (Atmaja, 2020)

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Document Accepted 13/9/24

2.5. Text Mining

Text Mining ialah proses ekstraksi informasi dari sumber data yang belum terstruktur, yang mencakup teknik penambangan data untuk analisis dan pemrosesan data. Dalam konteks text mining, langkah pertama melibatkan pengambilan data, yang kemudian mengalami tahap preprocessing sebelum dilakukan proses klasifikasi. Proses preprocessing melibatkan beberapa tahap, termasuk case folding, tokenization, filtering, dan penghapusan stop words. (Duei Putri dkk., 2022)

Teknik *Text Mining* dan model *decision tree* dapat digunakan untuk melakukan klasifikasi sebagai landasan dalam mengembangkan sistem pendukung keputusan untuk menentukan penugasan penanganan keluhan kepada divisi tertentu. (Galih Pradana, 2020)

2.6. Algoritma *Adaboost*

Algoritma *Adaboost* adalah suatu metode yang menghasilkan pohon keputusan yang lebih simpel dan menggunakan memori dengan lebih efisien. Algoritma *Adaboost* dapat melakukan klasifikasi pada model prediksi yang berbentuk pohon *(tree)* dan aturan berbasis aturan *(rule-based)*. (Gultom, 2020)

Adaboost merupakan suatu teknik dalam bidang data mining yang dapat meningkatkan tingkat akurasi metode klasifikasi. Teknik Adaboost dapat dengan mudah disatukan dengan metode klasifikasi dalam data mining, dengan menggunakan *feature selection* untuk memilih fungsi klasifikasi yang lemah, yang kemudian digabungkan bersama untuk membentuk fungsi klasifikasi yang baru. (Novianti dkk., 2022)

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

1. Input : Suatu kumpulan sampel penelitian dengan label $\{(\boldsymbol{x_i}, \boldsymbol{y_i}), \ldots, (\boldsymbol{x_N}, \boldsymbol{y_N})\}$ *Suatu Component Learn* Algoritma, jumlah perputaran T.
2. Initialize : Bobot suatu sampel pelatihan $\boldsymbol{w_i^1} = \frac{\boldsymbol{1}}{\boldsymbol{n}}$, untuk semua $\boldsymbol{i}$ = 1…, N
3. Lakukan untuk t =1, …., T
4. Gunakan *component learn* algoritma untuk melatih suatu klasifikasi, pada sampel bobot pelatihan
5. Hitung kesalahan pelatihannya pada $\boldsymbol{h_t} : \boldsymbol{\varepsilon_t} = \boldsymbol{\Sigma_i^n w_i^t}, \boldsymbol{y_i} \neq \boldsymbol{h_t(x_i)}$
6. Tetapkan bobot untuk *component clasifier* $\boldsymbol{h_t} = \boldsymbol{a_t} = \frac{\boldsymbol{1}}{\boldsymbol{2}} \boldsymbol{\ln} \left(\frac{\boldsymbol{1-\varepsilon_t}}{\boldsymbol{\varepsilon_t}}\right)$
7. Update bobot sampel pelatihan $\boldsymbol{w_i^{t+1}} = \frac{\boldsymbol{w_i^t \exp\{-a_t y_{iht(x_i)}\}}}{\boldsymbol{C_t}}, \boldsymbol{i = 1 \ldots, N}$ $\boldsymbol{C_t}$ adalah suatu konstanta normalisasi
8. Output $\boldsymbol{f(x) = sign(\Sigma_{t=1}^T a_t h_t(x))}$ (Qadrini L dkk., 2021)

2.7. Python

Python adalah bahasa pemrograman tingkat tinggi yang mendukung pendekatan pemrograman berorientasi objek. *Python* membedakan dirinya dari bahasa pemrograman lain melalui gaya penulisan sintaksnya yang khas. Dalam konteks bahasa pemrograman *Python*, terdapat berbagai perpustakaan (*library*) dan kerangka kerja (*framework*) yang digunakan untuk melakukan analisis data. (Duei Putri dkk., 2022)

Bahasa pemrograman Python menjadi pilihan yang populer dengan beberapa keunggulan berikut:

Document Accepted 13/9/24
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

1. Memudahkan penggunaan dalam pengembangan produk perangkat lunak, perangkat keras, *Internet of Things*, aplikasi *web*, dan *video game*.
2. Tidak hanya memiliki tingkat keterbacaan kode yang tinggi untuk mempermudah pemahaman, Python juga dilengkapi dengan beragam dan luasnya perpustakaan (*library*).
3. Merupakan bahasa yang sangat mendukung ekosistem *Internet of Things*. (Kadarina & Ibnu Fajar, 2019)

2.8. Confusion Matrix

Confusion matrix adalah teknik evaluasi yang digunakan untuk mengukur akurasi, menghitung nilai presisi berdasarkan positif, dan merangkum jumlah prediksi benar dan salah dalam tabel hitungan, yang kemudian dibagi per kelas. Penerapan *confusion matrix* juga memungkinkan visualisasi performa algoritma, terutama pada algoritma yang termasuk dalam pembelajaran terawasi. Informasi yang diberikan oleh confusion matrix tidak hanya menyoroti kesalahan dalam klasifikasi, melainkan juga memberikan wawasan mengenai kesalahan yang terjadi.

Rumus untuk masing masing elemen dalam Confusion Matrix adalah:

1. Presisi *(Precision*)

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FP}}$$

Precision mengukur proporsi dari hasil prediksi positif yang benar dari semua hasil prediksi positif.

2. Sensitivitas atau *True Positive Rate* (*Recall*)

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FN}}$$

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

Recall mengukur proporsi dari contoh positif yang benar diprediksi oleh model.

3. *Specificity (True Negative Rate):*

$$\boldsymbol{Specificity = \frac{TN}{TN + FN}}$$

Specificity mengukur proporsi dari contoh negatif yang benar diprediksi oleh model.

4. F1-*Score*

$$\boldsymbol{F1\ Score = 2 * \frac{Precision * Recall}{Precision * Recall}}$$

F1-Score adalah ukuran gabungan dari Precision dan Recall yang memberikan nilai tunggal yang menggambarkan keseimbangan antara keduanya.

True Positive (TP): Jumlah contoh positif yang benar-benar diprediksi dengan benar oleh model.

True Negative (TN): Jumlah contoh negatif yang benar-benar diprediksi dengan benar oleh model.

False Positive (FP): Jumlah contoh negatif yang salah diprediksi sebagai positif oleh model.

False Negative (FN): Jumlah contoh positif yang salah diprediksi sebagai negatif oleh model.

Confusion Matrix akan tampak seperti ini:

Tabel 2. 1 Confusion Matrix

	Predicted Positive	*Predicted Negative*
Actual Positive	*True Positive (TP)*	*False Negative(FN)*

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

Actual Negative	*False Positve(FN)*	*True Negative(TN)*

Dalam format tabel, confusion matrix menggambarkan prestasi dan ketidakberhasilan model klasifikasi dalam mengantisipasi kelas. Dengan menyelidiki aspek-aspek ini, kita dapat melakukan evaluasi yang lebih terperinci terhadap performa model. (Tatara dkk., 2023)

2.9. *Term Frequncy Inverse Document Frequency* (TF-IDF)

Pemberian bobot pada setiap kata, dikenal sebagai *term weighting*, merupakan langkah kunci dalam meningkatkan kapabilitas analisis sentimen dalam ranah *text mining*. Studi ini mengadopsi pendekatan *Term Frequency-Inverse Document Frequency (TF-IDF)*. *Term Frequency* (tf(w,d)) diartikan sebagai nilai yang mencerminkan signifikansi kata berdasarkan total kemunculannya dalam suatu teks atau dokumen. Sementara itu, *Inverse Document Frequency (IDF)* berperan sebagai metode pembobotan token yang memantau seberapa sering token muncul dalam suatu kumpulan teks. (Gifari dkk., 2022)

Pendekatan *TF-IDF (Term Frequency Inverse Document Frequency)* menciptakan suatu metode unik untuk mengatributkan bobot yang menggambarkan hubungan antara suatu kata (*term*) dan dokumen. Pendekatan ini merangkum dua konsep utama untuk perhitungan bobot, yaitu frekuensi kemunculan suatu kata dalam konteks tertentu dan invers frekuensi dokumen yang mengandung kata tersebut. Pada perhitungan *Term Frequency* (tf), nilai tf ditetapkan berdasarkan seberapa sering kata tersebut muncul dalam dokumen. Sebagai ilustrasi, jika kata tersebut muncul sebanyak lima kali, nilai tf-nya akan menjadi lima. Sementara itu,

Document Accepted 13/9/24
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Inverse Document Frequency (idf) dihitung melalui formula yang mencerminkan *invers* dari frekuensi dokumen yang mengandung kata tersebut.:

1. *Term Frequency* (TF)

$$TF(t,d) = \frac{jumlah\ kata\ t\ dalam\ dokumen\ d}{total\ kata\ dalam\ dokumen\ d}$$

Di mana:

t adalah kata yang ingin dihitung TF-IDF-nya.

d adalah dokumen di mana kata tersebut dihitung.

2. TF Normalisasi

Pengnormalan *Term Frequency* (TF) merupakan tahapan normalisasi bobot kata dalam suatu dokumen, bertujuan mengurangi dampak dominasi kata-kata yang muncul secara berlebihan di dokumen tertentu, sehingga tetap memperhatikan representasi relatif kata-kata dalam dokumen tersebut. Salah satu pendekatan umum untuk pengnormalan *Term Frequency* adalah menggunakan metode Normalisasi Logarithmic: $TF\ Normalized = 1 + \log(TF)$

Di mana :

TF-Normalized merujuk pada nilai normalisasi dari *Term Frequency*. *TF* sendiri mencerminkan nilai *Term Frequency* sebuah kata dalam suatu dokumen. Pendekatan Normalisasi *Logarithmic* ini bertujuan mengurangi efek perbedaan yang signifikan dalam frekuensi kata-kata. Dengan memanfaatkan logaritma, nilai *TF* yang tinggi akan disesuaikan ke tingkat yang lebih rendah, sehingga mengurangi disparitas antara kata-kata yang muncul sangat sering dan yang jarang.

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

Normalisasi *TF* ini tidak mengubah urutan relatif dari frekuensi kata-kata dalam dokumen, tetapi membantu dalam mengurangi efek dari frekuensi kata yang berlebihan, sehingga membantu dalam meningkatkan representasi kata-kata yang lebih seimbang dalam dokumen. Metode normalisasi lainnya juga dapat digunakan bergantung pada kebutuhan spesifik dalam konteks pemrosesan teks.

3. *Inverse Document Frequency* (IDF)

$$IDF(t) = \log\left(\frac{total\ dokumen}{jumlah\ dokumen\ yang\ mengandung\ kata\ t}\right)$$

Di mana:

t adalah kata yang ingin dihitung IDF-nya

Total dokumen adalah jumlah keseluruhan dokumen dalam korpus teks

Jumlah dokumen yang mengandung kata t adalah jumlah dokumen dimana kata t muncul.

4. TF-IDF

$$TF - IDF(t, d) = TF(t, d) * IDF(t)$$

Ini adalah hasil dari perkalian antara *Term Frequency (TF) dan Inverse Document Frequency (IDF)* dari sebuah kata dalam suatu dokumen.

TF-IDF memberikan penilaian yang lebih tinggi pada kata-kata yang muncul secara umum di suatu dokumen tetapi kurang umum di dokumen lain dalam kumpulan data, membantu dengan penekanan pada kata-kata yang bersifat unik atau signifikan dalam dokumen tersebut. Pendekatan ini digunakan secara luas dalam aplikasi beragam, termasuk sistem pencarian informasi, pengelompokan dokumen, dan analisis teks. (Gifari dkk., 2022)

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

2.10. Penelitian Terdahulu

Berikut ini adalah contoh penelitian terdahulu mengenai klasifikasi topik berita dan analisis sentimen menggunakan Adaboost :

Tabel 2. 2 Penelitian Terdahulu

No	Nama	Judul	Metode	Hasil
1.	Laurentius Dandi Andika, Dela Regita Cahyani, Dhika Saputra, Tirta Herawati, Muhammad Khoiruddinsyah, Dedi Dwi Saputra	Analisis Sentimen Konsumen KFC Berdasarkan Pendekatan Naïve Bayes dan Adaboost Berbasis Data Twitter	Naïve Bayes dan Adaboost	Dalam eksperimen ini, terdapat sebanyak 2214 data yang mencakup sentimen positif, negatif, dan netral. Dengan penerapan algoritma Naïve Bayes, hasilnya menunjukkan tingkat akurasi sebesar 73,04%, presisi mencapai 65,28%, dan recall mencapai 98,61%. Sementara itu, penggunaan algoritma Adaboost juga menghasilkan performa yang serupa, dengan akurasi, presisi, dan recall masing-masing mencapai 73,04%, 65,28%, dan 98,61%.
2.	Lila Dini Utami dan Romi Satria Wahono	Integrasi Metode Information Gain Untuk Seleksi Fitur dan Adaboost Untuk Mengurangi Bias pada Analisis Sentimen Review Restoran	Naïve Bayes dan Adaboost	Dalam penelitian ini, digunakan 200 data ulasan restoran, terbagi menjadi 100 ulasan positif dan 100 ulasan negatif. Performa algoritma Naïve Bayes menunjukkan akurasi sebesar 70%, sementara Adaboost menunjukkan hasil yang luar biasa dengan

Document Accepted 13/9/24
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

		Menggunakan Algoritma Naïve Bayes.		akurasi mencapai 99,50%.
3.	Hartati, Deni Hermawan, M. Akhsanal, Zailani Wahyudi, Angga Ariyanto, Dedi Dwi Saputra	Optimasi Analisis Sentimen Pada Olshop Tokopedia Menggunakan Text Mining Dengan Algoritma Naïve Bayes & Adaboost.	Naïve Bayes dan Adaboost	Dalam riset ini, diterapkan data berbahasa Indonesia yang berasal dari platform Twitter sejumlah 1000, terklasifikasi menjadi 941 positif dan 59 negatif. Dalam uji coba, Adaboost dan Naïve Bayes menunjukkan performa yang sangat memuaskan dengan mencapai tingkat akurasi 94,95%, presisi sebesar 90,86%, dan recall mencapai 100%.
4.	Aditya Wahyu Nur Faizi dan Kristiawan Nugroho	Penerapan Metode Adaptive Boosting Pada Analisis Sentimen Kenaikan BBM Pertamina.	Adaptive Boosting	Dalam penelitian ini, data diambil langsung dari tweet dengan total 560, yang kemudian dikelompokkan menjadi kategori positif dan negatif. Tingkat akurasi yang tercapai dalam penelitian ini mencapai 86,8%.
5.	Kurnia, Indah Purnamasari, dan Dedi Dwi Saputra	Analisis Sentimen dengan Metode Naïve Bayes, SMOTE, dan Adaboost pada Twitter Bank BTN.	Naïve Bayes, SMOTE, dan Adaboost	Penelitian ini menghimpun data dari akun Twitter Bank BTN dengan jumlah total 500 rekaman pada rentang waktu Desember 2021 hingga Maret 2022. Data tersebut kemudian dibagi

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

				menjadi dua kelompok, yaitu sentimen positif dan negatif. Dalam evaluasi menggunakan kombinasi Adaboost + Naïve Bayes, tercatat hasil akurasi sebesar 77,80%, presisi mencapai 83,27%, dan recall sebesar 83,91%. Sementara itu, setelah mengimplementasikan algoritma SMOTE + Adaboost + Naïve Bayes, terdapat peningkatan signifikan dengan akurasi mencapai 87,05%, presisi 90,63%, dan recall 83%

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

BAB III

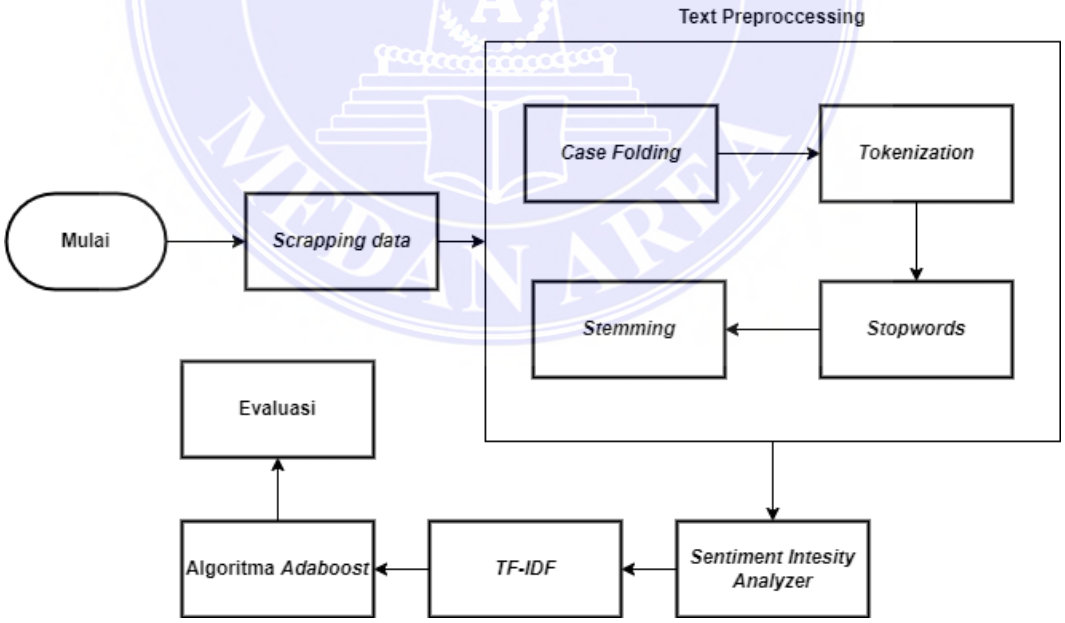
METODE PENELTIAN

3.1. Waktu dan Tempat Penelitian

Waktu yang digunakan pada pelaksanaan penelitian ini adalah semester genap tahun ajaran 2022/2023. Pada penelitian ini yang menjadi obyek penelitian adalah Twitter dengan *query* '*Seagames eSports Cambodia*'. Untuk data yang diambil, rentang waktunya adalah pada bulan November 2019 – Agustus 2023.

3.2. Metodologi Penelitian

Pada tahapan ini akan dijelaskan bagaimana tahapan-tahapan alur sistem mulai dari proses-proses pengambilan data sampai *output* yang dihasilkan ketika menggunakan Algoritma *Adaboost.*



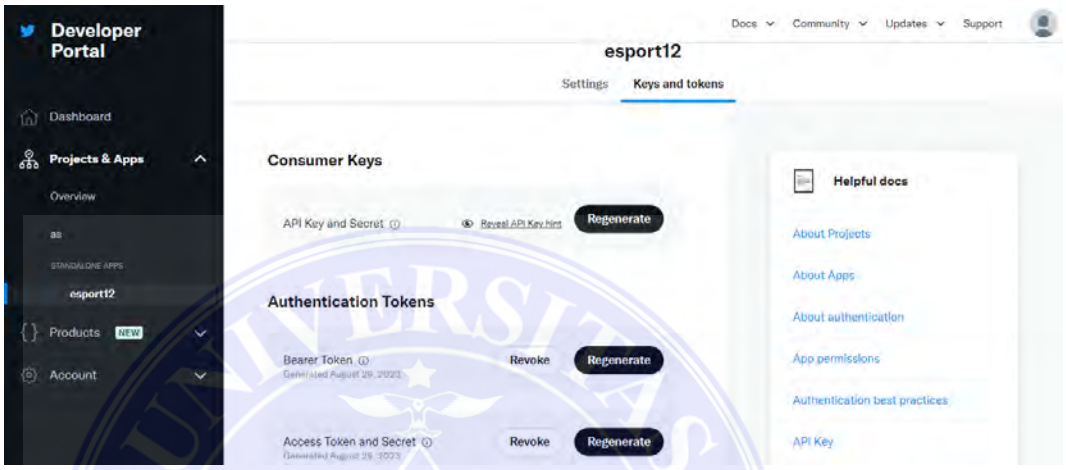
Gambar 3. 1. Alur Proses Penelitian

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Document Accepted 13/9/24

3.2.1. Pengumpulan data (*Scrapping Data*)

Penelitian ini telah mengumpulkan data sebanyak 1153 *record* dengan menggunakan API (*Application Programming Interface*) *Twitter*, yang mana berasal dari website *developer.twiiter.com* dapat dilihat pada gambar 3.2



Gambar 3. 2. Portal developer twitter

Untuk mengaksesnya dibutuhkan sebuah key yang berasal dari Twitter dengan cara registrasi ke portal developer Twitter yang berfungsi untuk mengkonfirmasi kepada pihak Twitter dalam memberikan ijin menjelajahi terkait data yang berkaitan dengan Twitter. Kemudian akan didapatkan key diantaranya:

API_key=’YAMIW6IrMCtTwBg5LLZd6KCSZ’,APIkeysecret=’wlW4Wc0Brt2WwFIRZfCmYUDc4WhGYHS7pHLuzF8TbHib0mGgVi’, API key adalah sebuah kunci atau token yang digunakan untuk mengidentifikasi dan mengotentikasi permintaan yang dibuat ke sebuah aplikasi atau layanan yang menyediakan Application Programming Interface (API). API key ini diberikan kepada pengguna atau pengembang sebagai cara untuk mengizinkan akses terhadap API tersebut, serta untuk melacak penggunaan dan mengatur hak akses. Dengan menggunakan API key, penyedia layanan

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

dapat memastikan bahwa hanya pengguna yang sah yang dapat mengakses dan menggunakan fungsionalitas yang disediakan oleh API mereka. accesstoken='1303155258aI3v3TH2VfwJvyKFmVFFpm9dF96T8kTWS2Zk4O6',accesstokensecret='c5up7cNGt0fAq1Giluzad3rpYptrGm4ThYVu9zMwxJguP'. Access token adalah sebuah token atau tanda yang diberikan kepada pengguna setelah mereka berhasil melakukan autentikasi dan otorisasi terhadap sebuah sistem atau layanan. Access token digunakan untuk memberikan izin akses terhadap sumber daya yang dilindungi oleh sistem tersebut, seperti data atau fungsi tertentu dalam API. Key tersebut yang akan digunakan untuk proses integrase antara Twitter API dengan bahasa pemrograman Python.

```python
# Scraping data menggunakan API Twitter
import tweepy # library
import pandas as pd # library


# Menggunakan token dari twitter
api_key= 'YAMIW6IrMCtTwBg5LLZd6KCSZ'
api_key_secret= 'wlW4Wc0Brt2WwFIRZfCmYUDc4WhGYHS7pHLuzF8TbHib0mGgVi'
access_token='1303155258-aI3v3TH2VfwJvyKFmVFFpm9dF96T8kTWS2Zk4O6'
access_token_secret='c5up7cNGt0fAq1Giluzad3rpYptrGm4ThYVu9zMwxJguP'

auth = tweepy.OAuthHandler(api_key, api_key_secret)
auth.set_access_token(access_token, access_token_secret)

api = tweepy.API(auth, wait_on_rate_limit=True)

# menentukan kata kunci
search_words = "sea games esport -filter:retweets"

tweets = tweepy.Cursor(api.search_tweets, q = search_words, count = 1153, include_entities = True).items(1153)

users_locs = [[tweet.user.username, tweet.text] for tweet in tweets]

tweet_text = pd.DataFrame(data = users_locs, columns = ["nama_pengguna", "teks"]) # nama kolom & teks

tweet_text.to_csv('seagames.xlsx', index=False) # simpan menjadi file excel
```

Gambar 3. 3. Scrape Data Menggunakan API

Pada Gambar 3.3 adalah proses pengambilan data menggunakan API Twitter yang mana dibutuhkan API_key, API_key_secret, access_token, access_token_secret. Adapun kata kunci yang dipakai ialah sea games esport

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24

Access From (repository.uma.ac.id)13/9/24

dan variabel yang akan diambil dari Twitter yaitu id, nama pengguna dan teks. Library yang digunakan pada python adalah tweepy dan pandas. Setelah data sudah didapatkan berbentuk file excel.

3.2.2. *Text Preproccessing*

Text preprocessing adalah serangkaian langkah atau proses yang dilakukan pada teks mentah (input) sebelum analisis atau pemrosesan lebih lanjut. Tujuan utama dari text preprocessing adalah untuk membersihkan, mengubah format, dan mengatur teks agar lebih mudah dipahami dan diolah oleh algoritma atau model tertentu. Pada penelitian tahapan text preprocessing yaitu case folding, tokenization, stopwords, dan stemming. Pada Proses ini dan seterusnya 3 data sampel yang diambil melalui *crawling* data *twitter* akan digunakan sebagai contoh sebagai berikut :

Tabel 3. 1 Data Sampel

Nama Pengguna	Teks
verov2929	GO GO AYO SEMANGAT PAK IBNU RIZA MENGEMBAN POSISI DI MENPORA, untuk masalah background mah ga perlu ragu lagi guys apalagi pernah menjabat sbg Ketua Badan Timnas Esport Sea Games 2021
giriramdan	wong kemaren aja tim aov indo gak jadi otw ikut sea games 2021 dengan dalih gak berprestasi dll...1soal begitu aja banyak alasan gimana mau bikin turnamen esport gede

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

jojopatricka	yauda mau kawal esport PUBG sea games dulu.

a) *Case Folding*

Pada proses ini semua karakter teks dalam suatu dokumen atau data ke dalam huruf kecil (lowercase).

Tabel 3. 2 Text to Case Folding

Teks	Case Folding
GO GO AYO SEMANGAT PAK IBNU RIZA MENGEMBAN POSISI DI MENPORA, untuk masalah background mah ga perlu ragu lagi guys apalagi pernah menjabat sbg Ketua Badan Timnas Esport Sea Games 2021	go go ayo semangat pak ibnu riza mengemban posisi di menpora, untuk masalah background mah ga perlu ragu lagi guys apalagi pernah menjabat sbg ketua badan timnas esport sea games
wong kemaren aja tim aov indo gak jadi otw ikut sea games 2021 dengan dalih gak berprestasi dll...1soal begitu aja banyak alasan gimana mau bikin turnamen esport gede	wong kemaren aja tim aov indo gak jadi otw ikut sea games dengan dalih gak berprestasi dll soal begitu aja banyak alasan gimana mau bikin turnamen esport gede
yauda mau kawal esport PUBG sea games dulu.	yauda mau kawal esport pubg sea games dulu.

b) *Tokenization*

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

Pada tahap ini *tokenization* digunakan untuk memecah teks menjadi kata-kata atau frasa-frasa, yang disebut "token." Token-token ini nantinya dapat dihitung, dianalisis, atau digunakan dalam berbagai tugas NLP seperti analisis sentimen.

Tabel 3. 3 Case Folding to Tokenization

Case Folding	Tokenization
go go ayo semangat pak ibnu riza mengemban posisi di menpora, untuk masalah background mah ga perlu ragu lagi guys apalagi pernah menjabat sbg ketua badan timnas esport sea games 2021	[“go”], [“go”], [“ayo”], [“semangat”], [“pak”], [“ibnu”], [“riza”], [“mengemban”], [“posisi”], [“di”], [“menpora”], [“untuk”], [“masalah”], [“background”], [“mah”], [“ga”], [“perlu”], [“ragu”], [“lagi”], [“guys”], [“apalagi”], [“pernah”], [“menjabat”], [“sbg”], [“ketua”], [“badan”], [“timnas”], [“esport’], [“sea,games”]
wong kemaren aja tim aov indo gak jadi otw ikut sea games 2021 dengan dalih gak berprestasi dll...1soal begitu aja banyak alasan gimana mau bikin turnamen esport gede	[“wong”], [“kemaren”], [“aja”], [“tim”], [“aov”], [“indo”], [“gak”], [“jadi”], [“otw”], [“ikut], [“sea”], [“games”], [“dengan”], [“dalih”], [“gak”], [“berprestasi”], [“dll”], [“soal”], [“begitu”], [“aja”], [“banyak”], [“alasan”], [“gimana”],

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

	[“mau”], [“bikin”], [“turnamen”], [“esport”], [“gede”]
yauda mau kawal esport pubg sea games dulu.	[“yauda”], [“mau”], [“kawal”], [“esport”], [“pubg”], [“sea”], [“games”], [“dulu”]

c) *Stopwords*

Pada tahap ini Stopwords menghapus kata-kata yang dianggap umum, sering muncul, dan memiliki sedikit nilai informasi dalam analisis teks.

Tabel 3. 4 Tokenization to Stopwords

Tokenization	Stopwords
[“go”], [“go”], [“ayo”], [“semangat”], [“pak”], [“ibnu”], [“riza”], [“mengemban”], [“posisi”], [“di”], [“menpora”], [“untuk”], [“masalah”], [“background”], [“mah”], [“ga”], [“perlu”], [“ragu”], [“lagi”], [“guys”], [“apalagi”], [“pernah”], [“menjabat”], [“sbg”], [“ketua”], [“badan”], [“timnas”], [“esport’], [“sea”], [“games”]	[“go”], [“go”], [“ayo”], [“semangat”], [“ibnu”], [“riza”], [“mengemban”], [“posisi”], [“menpora”], [“background”], [“mah”], [“ga”], [“ragu”], [“guys”], [“menjabat”], [“sbg”], [“ketua”], [“badan”], [“timnas”], [“esport’], [“sea”], [“games”]
[“wong”], [“kemaren”], [“aja”], [“tim”], [“aov”], [“indo”], [“gak”],	[“kemaren”], [“aja”], [“tim”], [“aov”], [“indo”], [“gak”], [“otw”],

Document Accepted 13/9/24
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

[“jadi”], [“otw”], [“ikut], [“sea”], [“games”], [“dengan”], [“dalih”], [“gak”], [“berprestasi”], [“dll”], [“soal”], [“begitu”], [“aja”], [“banyak”], [“alasan”], [“gimana”], [“mau”], [“bikin”], [“turnamen”], [“esport”], [“gede”]	[“sea”], [“games”], [“dalih”], [“gak”], [“berprestasi”], [“dll”], [“aja”], [“alasan”], [“gimana”], [“bikin”], [“turnamen”], [“esport”], [“gede”]
[“yauda”], [“mau”], [“kawal”], [“esport”], [“pubg”], [“sea”], [“games”], [“dulu”]	[“yauda”], [“kawal”], [“esport”], [“pubg”], [“sea”], [“games”],

Kata- kata yang dihapus antara lain “Pak”, “di”, “untuk”, “masalah”, “perlu”, “lagi”, “apalagi”, “pernah”, “nih”, “mau”, “wong”, “jadi”, “dengan”, “soal”, “begitu”, “banyak”, “mau”.

d) *Stemming*

Pada Tahap ini Stemming bertujuan untuk menghilangkan afiks (akhiran kata) dari kata-kata untuk menghasilkan bentuk dasar atau kata dasar (stem) yang lebih umum.

Tabel 3. 5 Stopwords to Steeming

Stopwords	Stemming
[“go”], [“go”], [“ayo”], [“semangat”], [“ibnu”], [“riza”], [“mengemban”], [“posisi”],	[“go”], [“go”], [“ayo”], [“semangat”], [“ibnu”], [“riza”], [“emban”], [“posisi”],

© Hak Cipta Di Lindungi Undang-Undang
Document Accepted 13/9/24
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

[“menpora”], [“background”], [“mah”], [“ga”], [“ragu”], [“guys”], [“menjabat”], [“sbg”], [“ketua”], [“badan”], [“timnas”], [“esport’], [“sea”], [“games”]	[“menpora”], [“background”], [“mah”], [“ga”], [“ragu”], [“guys”], [“jabat”], [“sbg”], [“ketua”], [“badan”], [“timnas”], [“esport’], [“sea”], [“games”]
[“kemaren”], [“aja”], [“tim”], [“aov”], [“indo”], [“gak”], [“otw”], [“sea”], [“games”], [“dalih”], [“gak”], [“berprestasi”], [“dll”], [“aja”], [“alasan”], [“gimana”], [“bikin”], [“turnamen”], [“esport”], [“gede”]	[“kemaren”], [“aja”], [“tim”], [“aov”], [“indo”], [“gak”], [“otw”], [“sea”], [“games”], [“dalih”], [“gak”], [“berprestasi”], [“dll”], [“aja”], [“alas”], [“gimana”], [“bikin”], [“turnamen”], [“esport”], [“gede”]
[“yauda”], [“kawal”], [“esport”], [“pubg”], [“sea”], [“games”],	[“yauda”], [“kawal”], [“esport”], [“pubg”], [“sea”], [“games”],

3.2.3. *Sentiment Intensity Analyzer*

Pada tahapan ini, Sentiment Intensity Analyzer adalah alat untuk menilai intensitas atau kekuatan sentimen dalam teks atau dokumen. Tujuan utama dari Sentiment Intensity Analyzer adalah untuk mengukur sejauh mana teks mengandung ekspresi perasaan atau sentimen, seperti positif atau negatif. Tahapan ini dibuat dengan menggunakan *Library* dengan bahasa pemrograman python.

Jika $\boldsymbol{x} \geq \mathbf{0}$ maka sentimen nya adalah positif,

Tabel 3. 6 Sentiment Intensity Aanlyzer

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

No	Negatif	Positif	*Compound* (x)	*Sentiment*
1	0,0	1,0	0,0	Positif
2	0,1	0,9	-0.2732	Negatif
3	0,0	1,0	0,0	Positif

3.2.4. *TF-IDF*

Pada tahap ini dilakukan pemrosesan teks dan pengambilan informasi untuk menilai pentingnya suatu kata atau frasa dalam dokumen teks atau koleksi dokumen (korpus). Metode ini digunakan untuk mengukur seberapa sering kata atau frasa tertentu muncul dalam dokumen tertentu (*Text Frequency* atau *TF*) sambil mempertimbangkan seberapa umum kata atau frasa tersebut dalam seluruh koleksi dokumen (*inverse document frequency* atau *IDF*).

Tabel 3. 7 TF-IDF

Term	TFNormalisasi			df	IDF	TF-IDF		
	D1	D2	D3			D1	D2	D3
Go	0,090	0	0	1	0,477	0,043407	0	0
Ayo	0,045	0	0	1	0,477	0,021465	0	0
Semangat	0,045	0	0	1	0,477	0,021465	0	0
Ibnu	0,045	0	0	1	0,477	0,021465	0	0
Riza	0,045	0	0	1	0,477	0,021465	0	0
Emban	0,045	0	0	1	0,477	0,021465	0	0
Posisi	0,045	0	0	1	0,477	0,021465	0	0
Menpora	0,045	0	0	1	0,477	0,021465	0	0
background	0,045	0	0	1	0,477	0,021465	0	0
Mah	0,045	0	0	1	0,477	0,021465	0	0
Ga	0,045	0	0	1	0,477	0,021465	0	0

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

Ragu	0,045	0	0	1	0,477	0,021465	0	0
Guys	0,045	0	0	1	0,477	0,021465	0	0
Jabat	0,045	0	0	1	0,477	0,021465	0	0
Sbg	0,045	0	0	1	0,477	0,021465	0	0
Ketua	0,045	0	0	1	0,477	0,021465	0	0
Badan	0,045	0	0	1	0,477	0,021465	0	0
Timnas	0,045	0	0	1	0,477	0,021465	0	0
Esport	0,045	0,05	0,167	3	0	0	0	0
Sea	0,045	0,05	0,167	3	0	0	0	0
Games	0,045	0,05	0,167	3	0	0	0	0
Kemaren	0	0,05	0	1	0,477	0	0,02385	0
Aja	0	0,1	0	1	0,477	0	0,0477	0
Tim	0	0,05	0	1	0,477	0	0,02385	0
Aov	0	0,05	0	1	0,477	0	0,02385	0
Indo	0	0,05	0	1	0,477	0	0,02385	0
Gak	0	0,1	0	1	0,477	0	0,0477	0
Otw	0	0,05	0	1	0,477	0	0,02385	0
Dalih	0	0,05	0	1	0,477	0	0,02385	0
Prestasi	0	0,05	0	1	0,477	0	0,02385	0
Dll	0	0,05	0	1	0,477	0	0,02385	0
Alas	0	0,05	0	1	0,477	0	0,02385	0
Gimana	0	0,05	0	1	0,477	0	0,02385	0
Bikin	0	0,05	0	1	0,477	0	0,02385	0
Turnamen	0	0,05	0	1	0,477	0	0,02385	0
Gede	0	0,05	0	1	0,477	0	0,02385	0
Yauda	0	0	0,167	1	0,477	0	0	0,079659
Kawal	0	0	0,167	1	0,477	0	0	0,079659
Pubg	0	0	0,167	1	0,477	0	0	0,079659

UNIVERSITAS MEDAN AREA

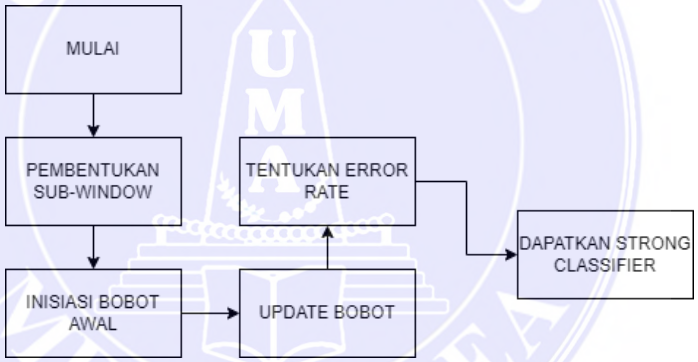
© Hak Cipta Di Lindungi Undang-Undang
---
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

3.3. Algoritma *Adaboost*

Adaboost bekerja dengan cara mengiterasi melalui dataset pelatihan dan memberikan bobot yang berbeda pada setiap sampel data. Pada setiap iterasi, bobot data yang salah diprediksi oleh model sebelumnya akan ditingkatkan, sehingga model berikutnya lebih fokus pada data yang sulit diprediksi. Proses ini berlanjut hingga sejumlah iterasi yang ditentukan sebelumnya atau sampai prediksi mencapai tingkat keakuratan yang diinginkan.

Pada penelitian ini, algoritma *Adaboost* akan di uji untuk mengklasifikasi sentimen negatif dan sentiment positif. berikut adalah diagram alur proses dari Algoritma *Adaboost*



Gambar 3. 4. Alur algoritma adaboost

3.4. Analisis Manual Data

Setelah mendapatkan nilai dari TF-IDF selanjutnya hasil dari TF-IDF D1, D2, dan D3 akan diproses menggunakan algoritma adaboost dan akan diprediksi sentimennya. Pada penelitian kali ini karena perbedaan jumlah kata yang ada pada data sampel dan keterbatasan jumlah kata pada dokumen 3, maka hasil yang akan digunakan adalah fitur 1,2 dan 3 (hasil TF-IDF kata ke-1, 2, dan 3). Hasil sentimen

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

aktual untuk positif adalah 1 dan negatif adalah -1. Berikut merupakan tabel dari data sampel yang akan digunakan untuk melakukan perhitungan manual.

Tabel 3. 8. Data sampel perhitungan manual

dokumen	Fitur-1	fitur-2	fitur-3	Sentimen aktual
d1	0,043	0,021	0,021	1
d2	0,023	0,047	0,023	-1
d3	0,079	0,079	0,079	1

Algoritma ini menggunakan *decision stump* (pohon keputusan dengan kedalaman 1) sebagai model lemah untuk adaboost. Inisialisasikan sampel data menjadi $\frac{1}{n}$, dimana n merupakan jumlah data. Dalam contoh ini $n = 3$. Selanjtunya akan diproses dengan beberapa langkah yaitu:

a) Inisialisasi Bobot : Setiap sampel data diberikan bobot yang sama yaitu $\frac{1}{3}$

b) Iterasi 1:

1. Latih model lemah *(Decision stump)* pada dataset pelatihan dengan bobot yang diberikan. Dari data sampel diatas, model lemah menghasilkan aturan: "Jika nilai Fitur-1,Fitur-2, dan Fitur-3 $\geq$ 0,04 maka prediksi positif (1); jika tidak negatif (-1)"

Tabel 3. 9. Data sampel setelah dilatih menggunakan algoritma adaboost

dokumen	Fitur-1	fitur-2	fitur-3	Sentimen aktual	sentimen prediksi
d1	0,043	0,021	0,021	1	-1
d2	0,023	0,047	0,023	-1	1
d3	0,079	0,079	0,079	1	1

*Label berwarna hijau merupakan hasil fitur yang memiliki nilai sentimen prediksi sesuai dengan sentimen aktual,

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

*Label berwarna kuning merupakan hasil fitur yang belum diubah dan belum diketahui sentimen prediksinya

*Label yang berwarna ungu merupakan nilai dari fitur yang nantinya akan diubah dengan menggunakan algortima adaboost

*Label berwarna merah merupakah label yang memiliki nilai sentimen yang salah dibandingkan dengan sentimen aktual

2. Evaluasi model lemah pada dataset pelatihan.

   Model lemah salah memprediksi dokumen 1 dan dokumen 2 pada fitur-2 dengan bobot kesalahan sebesar $\epsilon_1 = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$.

3. Hitung bobot model : $a_1 = \frac{1}{2}\ln\left(\frac{1-\epsilon_1}{\epsilon_1}\right) = \frac{1}{2}\ln\left(\frac{1-\frac{2}{3}}{\frac{2}{3}}\right) = -0{,}346$

4. Perbarui bobot sampel data :

   Dokumen 1 :$w_1 = \frac{1}{3} * e^{-0{,}346} = 0.235$

   Dokumen 2: :$w_2 = \frac{1}{3} * e^{-0{,}346} = 0.235$

Maka akan menghasilkan tabel prediksi sebagai berikut

Tabel 3. 10. Output algoritma adaboost

dokumen	Fitur-1	fitur-2	fitur-3	Sentimen aktual	sentimen prediksi
d1	0,043	0.235	0,021	1	1
d2	0,023	0.235	0,023	-1	1
d3	0,079	0,079	0,079	1	1

© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

Algoritma adaboost berhasil memprediksi d1 yang sebelumnya salah prediksi tetapi pada d2 hasil prediksi nya tetap salah dengan aturan awal pada iterasi 1 yaitu "Jika nilai Fitur-1,Fitur-2, dan Fitur-3 ≥ 0,04 maka prediksi positif (1); jika tidak negatif (-1)"

c) Gabungkan model : karena kita hanya melakukan satu kali literasi, model adaboost akan menggunakan hasil dari Iterasi 1

d) Output : model adaboost yang dihasilkan akan memprediksi sampel data berdasarkan hasil dari Iterasi 1

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil yang sudah dilakukan pada penelitian ini, dapat ditarik kesimpulan :

1. Hasil evaluasi metrik yang tertinggi yaitu pada pengujian ke-2 dengan data training 90% dan testing 10% mendapatkan akurasi sebesar 98%,
2. Hasil sentimen masyarakat terhadap Sea Games Esport yaitu positif dengan jumlah 111 (95,7%) dan hasil negatif berjumlah 5 (4,3%). Ini dapat disimpulkan bahwa masyarat twitter cenderung positif terhadap acara ttersebut.

5.2 Saran

Adapun saran yang dapat dilakukan terhadap penelitian selanjutnya yaitu :

1. Dapat menggunakan algoritma lain seperti Decision Tree, C.45, Random Forest, Multilayer Perceptron, dan Support Vector Machine.
2. Menambah jumlah data yang digunakan karena mempengaruhi hasil prediksi dan evaluasi metrik.
3. Dapat membuat sistem berbasis website menggunakan framework flask.

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Document Accepted 13/9/24

DAFTAR PUSTAKA

Adek Saputra, A. K. (2019). Study Minat Mahasiswa Angkatan Fakultas Ilmu Keolahrgaan Universitas Negeri Padang Terhadap Olahraga Cricket. *Progress in Retinal and Eye Research*, *561*(3), S2–S3.

Ahmed, S. T., Singh, D. K., Basha, S. M., Abouel Nasr, E., Kamrani, A. K., & Aboudaif, M. K. (2021). Neural Network Based Mental Depression Identification and Sentiments Classification Technique From Speech Signals: A COVID-19 Focused Pandemic Study. *Frontiers in Public Health*, *9*(December), 1–9. https://doi.org/10.3389/fpubh.2021.781827

Amrullah, A. Z., Sofyan Anas, A., & Hidayat, M. A. J. (2020). Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square. *Jurnal*, *2*(1), 40–44. https://doi.org/10.30812/bite.v2i1.804

Andhika, L. D., Cahyani, D. R., Saputra, D., Herawati, T., Khoiruddinsyah, M., & Saputra, D. D. (2023). Analisis Sentimen Kosumen KFC Berdasarkan Pendekatan Naive Bayes dan Ada Boost Berbasis Data Twitter. *Jurnal INSAN - Journal of Information System Management Innovation*, *3*(1), 55–61. https://doi.org/10.31294/jinsan.v3i1.2219

Atmaja, E. H. S. (2020). Prediksi Kemenangan eSport DOTA 2 Berdasarkan Data Pertandingan. *Avitec*, *2*(1), 31–38. https://doi.org/10.28989/avitec.v2i1.612

Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional. *Jurnal Tekno Kompak*, *15*(1), 131. https://doi.org/10.33365/jtk.v15i1.744

Dini, L., Sekolah, U., Informatika, T. M., Komputer, D., Mandiri, N., & Wahono, R. S. (2019). Integrasi Metode Information Gain Untuk Seleksi Fitur dan Adaboost Untuk Mengurangi Bias Pada Analisis Sentimen Review Restoran Menggunakan Algoritma Naïve Bayes. *Journal of Intelligent Systems*, *1*(2), 120–126.

Duei Putri, D., Nama, G. F., & Sulistiono, W. E. (2022). Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier. *Jurnal Informatika dan Teknik Elektro Terapan*, *10*(1), 34–40. https://doi.org/10.23960/jitet.v10i1.2262

Febriyani, E., & Februariyanti, H. (2022). Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Naive Bayes Di Twitter. *Jurnal TEKNO KOMPAK*, *17*(2), 25–38.

Galih Pradana, M. (2020). Penggunaan Fitur Wordcloud Dan Document Term Matrix Dalam Text Mining. *Jurnal Ilmiah Informatika (JIF)*, *8*(1), 38–43.

Gifari, O. I., Adha, M., Freddy, F., & Durrand, F. F. S. (2022). Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine. *Journal of Information Technology*, *2*(1), 36–40. https://doi.org/10.46229/jifotech.v2i1.330

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Document Accepted 13/9/24

Gultom, S. I. (2020). Implementasi Data Mining Menentukan Pola Hidup Sehat Bagi Pengguna KB Menggunakan Algoritma Adaboost (Studi Kasus :Dinas Serdang Bedagai). *Informasi dan Teknologi Ilmiah (INTI)*, *7*(3), 298–304. https://www.ejurnal.stmik-budidarma.ac.id/index.php/inti/article/view/2405

Hermawan, D., Akhsanal, M., Wahyudi, Z., Ariyanto, A., & Dwi, D. (2022). *Optimasi Analisis Sentimen Pada Twitter Olshop Tokopedia Menggunakan Textmining Dengan Algoritma Naïve Bayes & Adaboost*. *6*(September), 821–828. http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/493

Hidayatulloh, F. A., & Susanto, I. H. (2022). Analisis Kegagalan Bowling Timnas Indonesia Melawan Thailand Dalam Pertandingan Cricket T20I (Twenty 20 International) Di Sea Games Malaysia 2017. *Jurnal Kesehatan Olahraga*, *10*(03), 185–194. https://ejournal.unesa.ac.id/index.php/jurnal-kesehatan-olahraga/article/view/48313

Juliyanto, E., & Rusdi, F. (2019). Strategi Penyampaian Informasi Melalui Instagram Dengan Tampilan Infografis (di Kementerian Perdagangan RI). *Prologia*, *2*(2), 298. https://doi.org/10.24912/pr.v2i2.3591

Kadarina, T. M., & Ibnu Fajar, M. H. (2019). Pengenalan Bahasa Pemrograman Python Menggunakan Aplikasi Games Untuk Siswa/I Di Wilayah Kembangan Utara. *Jurnal Abdi Masyarakat (JAM)*, *5*(1), 11. https://doi.org/10.22441/jam.2019.v5.i1.003

Khudzaifah, K., Kristiyanto, A., Aprilijanto, T., & Riyadi, S. (2023). Analisis E Sport Sebagai Cabang Olahraga Baru. *Prosiding Simposium Nasional Multidisiplin (SinaMu)*, *4*, 416. https://doi.org/10.31000/sinamu.v4i1.7923

Kurnia, Purnamasari, I., & Saputra, D. D. (2023). Analisis Sentimen Dengan Metode Naïve Bayes, SMOTE Dan Adaboost Pada Twitter Bank BTN. *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, *7*(2), 235–242. https://doi.org/10.35870/jtik.v7i3.707

Lestari, R. W., Meiyuntariningsih, T., Ramadhani, H. S., & Psikologi, F. (2023). Kualitas tidur pada anggota komunitas olahraga elektronik: Bagaimana peran intensitas penggunaan game online? *INNER: Journal of Psychological Research*, *2*(4), 916–924.

Novaneliza, R., Handayani, F., Suhandar, R. J., Surono, H., Azzahra, N. S., & Nadilla, D. (2023). Perbandingan Algoritma Untuk Analisis Sentimen Pada Twitter Transportasi Umum Commuterline. *Jurnal Sains Komputer & Informatika (J-SAKTI*, *7*(1), 13–21.

Novianti, N., Zarlis, M., & Sihombing, P. (2022). Penerapan Algoritma Adaboost Untuk Peningkatan Kinerja Klasifikasi Data Mining Pada Imbalance Dataset Diabetes. *Jurnal Media Informatika Budidarma*, *6*(2), 1200. https://doi.org/10.30865/mib.v6i2.4017

Nugraha Hidayat, D., & Soenyoto, T. (2020). Cirebon Gaming Esport Club Management Article Info. *Journal of Physical Education and Sports*, *9*(1), 39–43. https://journal.unnes.ac.id/sju/index.php/jpes/article/view/36648

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Document Accepted 13/9/24

Qadrini L, Sepperwali A, & Aina A. (2021). Decision Tree dan Adaboostpada Klasifikasi Penerima Program Bantuan Sosial. *Decision Tree Dan Adaboost Pada Klasifikasi Penerima Program Bantuan Sosial*, *2*(7), 1959–1966.

Renjani, A. S., Irmansyah, J., Pendidikan, P., & Mandalika, U. P. (2023). Minat Mahasiswa Pendidikan Olahraga terhadap Cabang Olahraga E-Sports di Universitas Pendidikan Mandalika Sports Education Students ' Interest in E-Sports at Mandalika University of Education. *Empiricism Journal*, *4*(2), 629–640.

Rizal Ramli, M., & Sulastri, H. (2023). Sentiment Analysis Of Student Opinion Related To Online Learning Using Naïve Bayes Classifier Algorithm And SVM With Adaboost On Twitter Social Media Analisis Sentimen Terhadap Opini Mahasiswa Terkait Pembelajaran Daring Dengan Menggunakan Algoritma Naïve. *Jurnal Informatika dan Teknologi Informasi*, *20*(2), 187–201. https://doi.org/10.31515/telematika.v20i2.8827

Sari, F. V., & Wibowo, A. (2019). Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi. *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, *10*(2), 681–686. https://jurnal.umk.ac.id/index.php/simet/article/view/3487

Sejati, W., Singh Bist, A., & Tambunan, A. (2023). Pengembangan Analisis Sentimen dalam Rekayasa Software Engineering menggunakan tinjauan literatur sistematis. *Jurnal MENTARI : Manajemen Pendidikan dan Teknologi Informasi*, *2*(1), 95–103. https://journal.pandawan.id/mentari/article/view/377

Septian, J. A., Fachrudin, T. M., & Nugroho, A. (2019). Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor. *Journal of Intelligent System and Computation*, *1*(1), 43–49. https://doi.org/10.52985/insyst.v1i1.36

Tatara, B. A., Margiyati, T., & Timur, F. G. C. (2023). Ancaman Terorisme dan Strategi Penanganannya: Studi Pustaka Strategi Kontra-Terorisme di Kamboja. *JSHP : Jurnal Sosial Humaniora dan Pendidikan*, *7*(1), 91–100. https://doi.org/10.32487/jshp.v7i1.1638

Wahyu, A., Faizi, N., & Nugroho, K. (2023). *Penerapan Metode Adaptive Boosting Pada Analisis Sentimen Kenaikan BBM Pertamina*. *08*(2016), 171–180.

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Document Accepted 13/9/24

Lampiran

1. Source Code

```
# Scraping data menggunakan API Twitter
import tweepy # library
import pandas as pd # library


# Menggunakan token dari twitter
api_key= 'YAMIW6IrMCtTwBg5LLZd6KCSZ'
api_key_secret=
'wlW4Wc0Brt2WwFIRZfCmYUDc4WhGYHS7pHLuzF8TbHib0mGgVi'
access_token='1303155258-
aI3v3TH2VfwJvyKFmVFFpm9dF96T8kTWS2Zk4O6'
access_token_secret='c5up7cNGt0fAq1Giluzad3rpYptrGm4ThYVu9zMw
xJguP'

auth = tweepy.OAuthHandler(api_key, api_key_secret)
auth.set_access_token(access_token, access_token_secret)

api = tweepy.API(auth, wait_on_rate_limit=True)

# menentukan kata kunci
search_words = "sea games esport -filter:retweets"

tweets = tweepy.Cursor(api.search_tweets, q = search_words, count =
1153, include_entities = True).items(1153)

users_locs = [[tweet.user.username, tweet.text] for tweet in tweets]

tweet_text = pd.DataFrame(data = users_locs, columns =
["nama_pengguna", "teks"]) # nama kolom & teks

tweet_text.to_csv('seagames.xlsx', index=False) # simpan menjadi file
excel
# upload file seagames.xlsx

from google.colab import files

uploaded = files.upload()

print(uploaded)
# Read Dataset
import pandas as pd
```


Document Accepted 13/9/24
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

```
df = pd.read_excel('/content/seagames.xlsx')
df
# case folding

import re

def case_folding(teks):
  teks = re.sub(r'http\S+','',teks) # delete url
  teks = re.sub(r'@[A-Za-a0-9]+',' ',teks) # delete simbol dan angka
  teks = re.sub(r'#[A-Za-z0-9]+',' ',teks)
  teks = re.sub(r'[0-9]+',' ',teks) # delete angka
  teks = re.sub(r"[-()\"#/@;:<>{}'+=~|.!?,_]", " ", teks) # delete simbol
  teks = teks.strip(' ')
  teks = teks.lower() # mengubah menjadi lower
  return teks
df['case_folding']= df['teks'].apply(case_folding)
df.head(50)
# tokenisasi

import nltk
nltk.download('punkt')
from nltk.tokenize import word_tokenize # library

def tokenisasi(teks):
  teks = word_tokenize(teks)
  return teks
df['tokenization'] = df['case_folding'].apply(tokenisasi)
df.head(50)
# Stopwords

from nltk.corpus import stopwords # library
nltk.download('stopwords')

daftar_stopword = stopwords.words('indonesian') # setting bahasa
indonesia
daftar_stopword.extend(["yg","dg","lg","ni","kl","kat","trs","nih","dgn","
w"])
daftar_stopword = set(daftar_stopword)

def Stopwords(words):
 return [word for word in words if word not in daftar_stopword]

df['stopwords'] = df['tokenization'].apply(Stopwords)
df.head(50)
# instal terlebih dahulu

!pip install sastrawi
```

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

```
# Stemming

from Sastrawi.Stemmer.StemmerFactory import StemmerFactory # library
khusus mengatasi data yang berbahasa indonesia

factory = StemmerFactory()
stemmer = factory.create_stemmer()

def porterstemmer(text):
  text = ' '.join(stemmer.stem(word) for word in text if word in text)
  return text

df['stemming'] = df['stopwords'].apply(porterstemmer)
df.head(50)

import nltk
nltk.download('vader_lexicon')
from nltk.sentiment.vader import SentimentIntensityAnalyzer # Libraries

sentimen = SentimentIntensityAnalyzer()
polaritas = lambda x : sentimen.polarity_scores(x) # Kasih kondisi
menggunakan lambda
nilai_sentimen = df.stemming.apply(polaritas) # panggil kondisi lambda di
atas & apply ke dataset yang sudah dibersihkan
sentiment_data = pd.DataFrame(data = list(nilai_sentimen)) # ubah ke
dalam data frame (kolom) agar lebih rapi
sentiment_data.head(50)
# Menentukan Label positif dan negatif berdasarkan kolom compound (-1
sampai 1) di atas
# jika x >= 0 == 'positif' sisanya == 'negatif'

label = lambda x : 'Positif' if x>=0 else 'Negatif'
sentiment_data['Sentimen'] = sentiment_data.compound.apply(label)
sentiment_data.head(50)
# Gabungkan dataset dan Sentimennya

dataset = df.join(sentiment_data.Sentimen)
dataset.head() # default 5 data teratas

# Hasil jumlah sentimen otomatis (tidak menggunakan algoritma
adaboost)

dataset.Sentimen.value_counts()
# ubah kolom stemming menjadi string

stem = dataset['stemming'].astype(str)
```

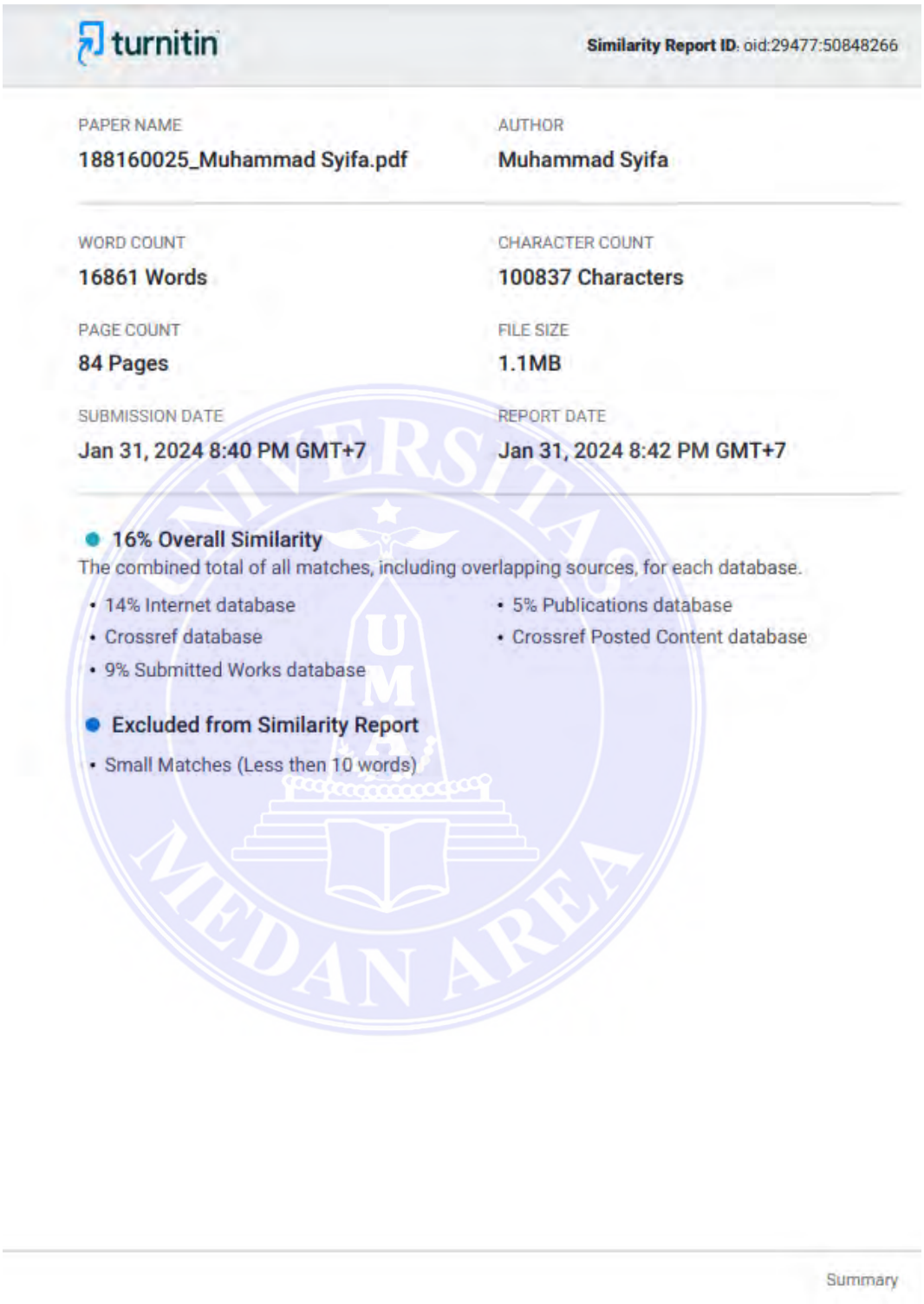

Document Accepted 13/9/24
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

```
from sklearn.feature_extraction.text import TfidfVectorizer # library TF-IDF

vectorizer = TfidfVectorizer()
vector = vectorizer.fit_transform(stem)
from sklearn.feature_extraction.text import TfidfVectorizer # library TF-IDF

vectorizer = TfidfVectorizer()
vector = vectorizer.fit_transform(stem)
from sklearn.model_selection import train_test_split # library membagi data training dan data testing

# Tentukan x & y
x = vector
y = dataset['Sentimen']

# split training dan testing
X_train, X_test, y_train, y_test = train_test_split(x, y, test_size = 0.1, random_state=0) # bisa pakai train_size atau test_size
from sklearn.ensemble import AdaBoostClassifier # library adabost

ab = AdaBoostClassifier(n_estimators=9000, random_state=0)

ab.fit(X_train, y_train) # melatih data pakai data training

y_pred = ab.predict(X_test) # prediksi pakai data testing
from sklearn.metrics import classification_report # library hasil evaluasi metrik: Akurasi

print(classification_report(y_test, y_pred))
# confusion matrix

from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay
import matplotlib.pyplot as plt

cm = confusion_matrix(y_true = y_test, y_pred = y_pred, labels=ab.classes_)
disp = ConfusionMatrixDisplay(confusion_matrix=cm, display_labels=ab.classes_)
disp.plot()

plt.show()
# membandingkan sentimen aktual terhadap sentimen prediksi

hasil = pd.DataFrame({'sentimen_aktual' : y_test, 'sentimen_prediksi' : y_pred}) hasil.head(50)
```

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

2. Turnitin



turnitin

Similarity Report ID: oid:29477:50848266

PAPER NAME
188160025_Muhammad Syifa.pdf

AUTHOR
Muhammad Syifa

WORD COUNT
16861 Words

CHARACTER COUNT
100837 Characters

PAGE COUNT
84 Pages

FILE SIZE
1.1MB

SUBMISSION DATE
Jan 31, 2024 8:40 PM GMT+7

REPORT DATE
Jan 31, 2024 8:42 PM GMT+7

16% Overall Similarity
The combined total of all matches, including overlapping sources, for each database.

- 14% Internet database
- 5% Publications database
- Crossref database
- Crossref Posted Content database
- 9% Submitted Works database

Excluded from Similarity Report

- Small Matches (Less then 10 words)

Summary

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

3. SK Pembimbing

UNIVERSITAS MEDAN AREA
FAKULTAS TEKNIK
Kampus I : Jalan Kolam Nomor 1 Medan Estate/Jalan PBSI Nomor 1 ☎ (061) 7366878, 7360168, 7364348, 7366781, Fax.(061) 7366998 Medan 20223
Kampus II : Jalan Setiabudi Nomor 79 / Jalan Sei Serayu Nomor 70 A, ☎ (061) 8225602, Fax. (061) 8226331 Medan 20122
Website: www.teknik.uma.ac.id E-mail: univ_medanarea@uma.ac.id

Nomor : 680/FT.6/01.10/IX/2023 — 20 September 2023
Lamp : -
H a l : **Perubahan Judul Tugas Akhir**

Yth, Pembimbing Tugas Akhir
Nurul Khairina, S. Kom, M. Kom
di
Tempat

Dengan hormat, Sehubungan dengan adanya perubahan judul tugas akhir maka perlu diterbitkan kembali SK Pembimbing Skripsi baru atas nama mahasiswa tersebut :

N a m a : Muhammad Syifa
N P M : 188160025
Jurusan : Teknik Informatika

Maka dengan hormat kami mengharapkan kesediaan saudara :

Nurul Khairina, S. Kom, M. Kom **(Sebagai Pembimbing)**

Adapun Tugas Akhir Skripsi berjudul :

"Analisis *Sentiment* Masyarakat terhadap *Sea Games E-Sport* pada *Twitter* dengan Menggunakan Algoritma *Adaboost*".

SK Pembimbing ini berlaku selama enam bulan terhitung sejak SK ini diterbitkan. Jika proses pembimbing melebihi batas waktu yang telah ditetapkan, SK ini dapat ditinjau ulang.

Demikian kami sampaikan, atas kesediaan saudara diucapkan terima kasih.

Dekan,

Dr. Rahmad Syah, S. Kom, M. Kom

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24

4. Pengantar Riset

UNIVERSITAS MEDAN AREA
FAKULTAS TEKNIK
Kampus I : Jalan Kolam Nomor 1 Medan Estate/Jalan PBSI Nomor 1 ☎ (061) 7366878, 7360168, 7364348, 7366781, Fax.(061) 7366998 Medan 20223
Kampus II : Jalan Setiabudi Nomor 79 / Jalan Sei Serayu Nomor 70 A, ☎ (061) 8225602, Fax. (061) 8226331 Medan 20122
Website: www.teknik.uma.ac.id E-mail: univ_medanarea@uma.ac.id

Nomor : 621 /FT.6/01.10/VIII/2023 11 Agustus 2023
Lamp : -
H a l : **Penelitian Dan Pengambilan Data Tugas Akhir**

Yth. Wakil Rektor Bid. Pengembangan SDM & Adm. Keuangan
Jln. Kolam No.1
Di
Medan

Dengan hormat, kami mohon kesediaan ibu kiranya berkenan untuk memberikan izin dan kesempatan kepada mahasiswa kami tersebut dibawah ini :

NO	NAMA	NPM	PRODI
1	Muhammad Syifa	188160025	Teknik Informatika

Untuk melaksanakan Penelitian dan Pengambilan Data Tugas Akhir di **Laboratorium Komputer Program Studi Teknik Informatika Fakultas Teknik Universitas Medan Area.**

Perlu kami jelaskan bahwa Pengambilan Data tersebut adalah semata-mata untuk tujuan Ilmiah dan Skripsi, yang merupakan salah satu syarat bagi mahasiswa tersebut untuk mengikuti ujian sarjana pada Fakultas Teknik Universitas Medan Area dan tidak untuk dipublikasikan, dengan judul :

Klasifikasi Topik Berita dan Analisis *Sentiment* pada *Tweets* Divisi Humas Polri dengan Menggunakan *Adaboost*.

Mohon kiranya tanggal Surat Izin Pengambilan Data Tugas Akhir agar disesuaikan dengan tanggal Terbitnya SK ini.

Atas perhatian dan kerja sama yang baik diucapkan terima kasih.

Dekan

Dr. Rahmad Syah, S. Kom, M. Kom

Tembusan :
1. Ka. BAMAI
2. Mahasiswa
3. File

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area
Access From (repository.uma.ac.id)13/9/24

Document Accepted 13/9/24

5. Selesai Riset

UNIVERSITAS MEDAN AREA
Kampus I : Jalan Kolam Nomor 1 Medan Estate ☎ (061) 7360168, 7366878, 7364348 🖷 (061) 7368012 Medan 20223
Kampus II : Jalan Setiabudi Nomor 79 / Jalan Sei Serayu Nomor 70 A ☎ (061) 8225602 🖷 (061) 8226331 Medan 20122
Website: www.uma.ac.id E-Mail: univ_medanarea@uma.ac.id

SURAT KETERANGAN
Nomor : 1888/UMA/B/01.7/X/2023

Rektor Universitas Medan Area dengan ini menerangkan bahwa :

Nama	: Muhammad svifa
No. Pokok Mahasiswa	: 188160025
Program Studi	: Teknik Informatika
Fakultas	: Teknik

Benar telah selesai Pengambilan Data Tugas Akhir di Laboratorium Komputer Universitas Medan Area dengan Judul Skripsi **"Klasifikasi Topik Berita dan Analisis Sentiment Pada Tweets Divisi Humas Polri Dengan Menggunakan Adaboost".**

Dan kami harapkan Data tersebut kiranya dapat membantu yang bersangkutan dalam penyusunan skripsi dan dapat bermanfaat bagi mahasiswa khususnya Fakultas Teknik

Demikian surat ini diterbitkan untuk dapat digunakan seperlunya.

Medan, 13 Oktober 2023.
a.n Rektor
Wakil Rektor Bidang Pengembangan SDM & Keuangan,

Hj. Dr. Ir. Suswati, MP

Tembusan :
1. Mahasiswa Ybs
2. File

UNIVERSITAS MEDAN AREA
© Hak Cipta Di Lindungi Undang-Undang
1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 13/9/24
Access From (repository.uma.ac.id)13/9/24