

SKRIPSI

**ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP
FASILITAS SEA GAMES 2023 MELALUI MEDIA TWITTER
MENGUNAKAN ALGORITMA *SUPPORT VECTOR
MACHINE***



AZKA LUTHFAN RUDIANA

1512619093

PROGRAM STUDI

PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER

FAKULTAS TEKNIK

UNIVERSITAS NEGERI JAKARTA

2024




**ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP FASILITAS SEA
GAMES 2023 MELALUI MEDIA TWITTER MENGGUNAKAN
ALGORITMA SUPPORT VECTOR MACHINE**

Azka Luthfan Rudiana, NIM: 1512619093

HALAMAN PENGESAHAN

NAMA DOSEN	TANDA TANGAN	TANGGAL
Dr. Widodo, S.Kom., M.Kom. Dosen Pembimbing I		19-07-2024
Ressy Dwitias Sari, S.T, M.T.I. Dosen Pembimbing II		19/07/2024

PENGESAHAN OLEH PANITIA UJIAN SKRIPSI

NAMA DOSEN	TANDA TANGAN	TANGGAL
M. Ficky Duskarnaen, M.Sc. Ketua Penguji		19/07/2024
Via Tuhamah Fauziastuti, M. Ed Dosen Penguji I		18 Juli 2024
Bambang Prasetya Adhi, S. Pd, M. Kom. Dosen Penguji II		19 Juli 2024

Tanggal Lulus: 19 Juli 2024

LEMBAR PERNYATAAN

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. skripsi ini berjudul "Analisis Sentimen Berbasis Aspek Terhadap Fasilitas Sea Games 2023 Melalui Media Twitter Menggunakan Algoritma *Support Vector Machine*" merupakan karya asli penulis dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta.
2. skripsi ini murni gagasan, rumusan, dan penelitian penulis sendiri dengan mendapatkan arahan dan bimbingan dari dosen pembimbing.
3. skripsi ini belum dipublikasikan, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpanan dan ketidakbenaran, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Universitas Negeri Jakarta.

Jakarta, 19 Juli 2024

Yang Membuat Pernyataan,



Azka Luthfan Rudiana

NIM. 1512619093



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS NEGERI JAKARTA
UPT PERPUSTAKAAN

Jalan Rawamangun Muka Jakarta 13220
Telepon/Faksimili: 021-4894221
Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Azka Luthfan Rudiana
NIM : 1512619093
Fakultas/Prodi : Pendidikan Teknik Informatika dan Komputer
Alamat email : azkaluthfanr@gamil.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada UPT Perpustakaan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

Skripsi Tesis Disertasi Lain-lain (.....)

yang berjudul :

ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP FASILITAS SEA GAMES

2023 MELALUI MEDIA TWITTER MENGGUNAKAN ALGORITMA SUPPORT

VECTOR MACHINE

Dengan Hak Bebas Royalti Non-Eksklusif ini UPT Perpustakaan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta

Penulis

(Azka Luthfan Rudiana)
nama dan tanda tangan

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat memulai dan menyelesaikan penyusunan skripsi dengan judul **“ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP FASILITAS SEA GAMES 2023 MELALUI MEDIA TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE”**, yang diajukan untuk memenuhi syarat kelulusan dalam meraih gelar Sarjana (S1) di Program Studi Pendidikan Teknik Informatika dan Komputer Fakultas Teknik Universitas Negeri Jakarta.

Penulis menyadari bahwa dalam proses penyusunan dan penulisan skripsi ini mengalami berbagai hambatan dan rintangan. Namun berkat doa, bimbingan, dukungan semangat serta saran dari berbagai pihak penulis dapat menyelesaikan penulisan skripsi ini. Maka dalam kesempatan ini, penulis dengan setulus hati menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT atas segala rahmat, karunia, hidayah, serta nikmat-Nya sehingga penulis dapat menyelesaikan skripsi ini.
2. Kedua orang tua yang penulis hormati dan sayangi, yaitu Bapak Moch, Agus Rudiana dan Ibu Eva Latifa yang senantiasa mendoakan yang terbaik untuk penulis dan selalu memberikan dukungan secara moril dan materiil kepada penulis.
3. Bapak Muhammad Ficky Duskarnaen, M.Sc., sebagai Koordinator Program Studi Pendidikan Teknik Informatika dan Komputer serta sebagai Dosen Pembimbing Akademik yang telah membimbing dan membantu penulis dengan berkonsultasi perihal akademik selama penulis berkuliah.
4. Bapak Dr. Widodo, S.Kom, M.Kom, sebagai Dosen Pembimbing I atas inspirasi, ide, arahan, bimbingan, serta ketersediaan waktunya selama penulisan dan penyusunan skripsi sehingga penulis dapat menyelesaikan skripsi dengan lancar.
5. Bu Ressy Dwitias Sari, S.T, M.T.I, sebagai Dosen Pembimbing II atas motivasi, inspirasi, arahan, bimbingan, serta ketersediaan waktunya selama penulisan dan penyusunan skripsi sehingga penulis dapat menyelesaikan skripsi dengan sebaik-baiknya.
6. Bapak/Ibu Dosen Program Studi Pendidikan Teknik Informatika dan Komputer atas ilmu yang telah diberikan selama perkuliahan yang secara tidak langsung memberikan dukungan dalam penulisan dan penyusunan skripsi.

7. Seluruh pihak yang terlibat secara langsung maupun tidak langsung yang tak bisa penulis sebutkan satu persatu.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna, maka dari itu penulis menyampaikan permohonan maaf apabila terdapat kekurangan dan kesalahan dalam proses penulisan skripsi ini, serta penulis mengharapkan kritik dan saran dari pembaca untuk penyempurnaan literatur skripsi ini. Penulis berharap agar penelitian dan penyusunan skripsi ini dapat bermanfaat bagi pembaca pada umumnya dan terkhususnya bagi penulis.

Jakarta, ... Juli 2024
Yang membuat pernyataan,

Azka Luthfan Rudiana
NIM. 1512619093



ABSTRAK

Azka Luthfan Rudiana, Penelitian yang berjudul “analisis sentimen berbasis aspek terhadap fasilitas sea games 2023 melalui media twitter menggunakan algoritma *support vector machine*” . Program Studi Pendidikan Teknik Informatika dan Komputer, Fakultas Teknik, Universitas Negeri Jakarta, 2024. Dosen Pembimbing : Dr. Widodo, M.Kom. dan Ressay Dwitias Sari, S.T, M.T.I.

Southeast Asian Games (SEA Games) merupakan festival multi-olahraga yang diadakan setiap dua tahun sekali di Asia Tenggara, dengan SEA Games ke-32 tahun 2023 diselenggarakan di Kamboja. Penyelenggaraan acara ini menjadi sorotan publik khususnya pada media sosial Twitter terkait kesiapan fasilitasnya, seperti lampu podium yang mati dan kamar atlet yang bocor, yang menimbulkan beragam tanggapan dari masyarakat. Analisis sentimen berbasis aspek terhadap fasilitas SEA Games 2023 ini menggunakan algoritma *Support Vector Machine* (SVM) untuk mengidentifikasi persepsi publik mengenai sentimen positif dan negatif serta aspek kebersihan, kelengkapan, dan kenyamanan dari fasilitas yang diberikan.

Data diambil dari tweet di Twitter, menghasilkan 555 data bersih. Model SVM diuji dengan *K-Fold Cross Validation* (K=10), menghasilkan akurasi 93,16%, presisi 90,94%, *recall* 83,96%, dan *F1-score* 85,58%. Performa model terbaik per fold terjadi pada fold ke-5 dengan akurasi 98,21%, *recall* 99%, dan *F1-score* 95,64%. Nilai presisi terbaik terjadi pada fold ke-6 sebesar 98,03%. Untuk aspek kenyamanan, akurasi mencapai 95,15%, presisi 90%, dan *F1-score* 80,81%, sedangkan nilai *recall* terbaik diperoleh dari aspek kebersihan sebesar 85%.

Kata Kunci: SEA Games, Aspek, *Support Vector Machine*, *K-Fold*, Twitter, Akurasi.

ABSTRACT

Azka Luthfan Rudiana, *Aspect-Based Sentiment Analysis of SEA Games 2023 Facilities through Twitter Using the Support Vector Machine Algorithm.* Department of Computer and Informatics Engineering Education, Faculty of Engineering, Universitas Negeri Jakarta, 2024. Advisors: Dr. Widodo, M.Kom. and Ressy Dwitias Sari, S.T., M.T.I.

The Southeast Asian Games (SEA Games) is a biennial multi-sport event held in Southeast Asia, with the 32nd SEA Games in 2023 hosted by Cambodia. The event's organization drew public scrutiny, particularly on social media platform Twitter, regarding the preparedness of its facilities, such as malfunctioning podium lights and leaking athlete accommodations, eliciting various responses from the public. This aspect-based sentiment analysis of SEA Games 2023 facilities employs the Support Vector Machine (SVM) algorithm to identify public perceptions of positive and negative sentiments and aspects of cleanliness, completeness, and comfort of the provided facilities.

The data was collected from tweets on Twitter, resulting in a clean dataset of 555 entries. The SVM model was tested using K-Fold Cross Validation ($K=10$), yielding an accuracy of 93.16%, a precision of 90.94%, a recall of 83.96%, and an F1-score of 85.58%. The best model performance per fold occurred in the 5th fold, with an accuracy of 98.21%, a recall of 99%, and an F1-score of 95.64%. The highest precision was achieved in the 6th fold, at 98.03%. For the comfort aspect, the accuracy reached 95.15%, precision was 90%, and F1-score was 80.81%, while the best recall value was obtained from the cleanliness aspect, at 85%.

Keywords: SEA Games, Aspects, Support Vector Machine, K-Fold, Twitter, Accuracy.

DAFTAR ISI

HALAMAN PENGESAHAN	I
LEMBAR PERNYATAAN.....	II
KATA PENGANTAR	IV
ABSTRAK.....	VI
ABSTRACT	VII
DAFTAR TABEL.....	X
DAFTAR GAMBAR.....	XI
DAFTAR LAMPIRAN	XII
1.1 LATAR BELAKANG.....	1
1.2 IDENTIFIKASI MASALAH	5
1.3 PEMBATAAN MASALAH.....	5
1.4 RUMUSAN MASALAH	6
1.5 TUJUAN PENELITIAN.....	6
1.6 MANFAAT PENELITIAN.....	6
BAB II KERANGKA TEORITIK	7
2.1 KERANGKA TEORITIK	7
2.1.1 SEA Games.....	7
2.1.2 Twitter	7
2.1.3 Analisis Sentimen.....	9
2.1.4 Analisis Sentimen Berbasis Aspek.....	9
2.1.5 Preprocessing	10
2.1.6 Support Vector Machine	13
2.1.7 K-Fold Cross Validation	15
2.1.8 Google Collab	16
2.1.9 Tweepy	16
2.2 PENELITIAN TERDAHULU	16
2.3 KERANGKA BERPIKIR.....	20
BAB III METODE PENELITIAN.....	23

3.1	TEMPAT DAN WAKTU.....	23
3.2	ALAT DAN BAHAN PENELITIAN	23
3.2.1	Alat.....	23
3.2.2	Bahan	24
3.3	DIAGRAM ALIR PENELITIAN.....	25
3.3.1	Melakukan <i>Crawling</i> untuk pengambilan data <i>tweet</i>	26
3.3.2	Mengidentifikasi Aspek pada data secara manual	26
3.3.3	Melabelkan data secara manual.....	26
3.3.4	<i>Preprocessing</i>	27
3.3.5	Menganalisis sentimen dengan SVM	30
3.3.6	Evaluasi hasil dengan K-Folds CV.....	30
3.4	TEKNIK DAN PROSEDUR PENGUMPULAN DATA.....	31
3.5	TEKNIK ANALISIS DATA	32
BAB IV HASIL PENELITIAN.....		33
4.1	PENGUMPULAN DATA	33
4.2	<i>PREPROCESSING DATA</i>	36
4.2.1	<i>Cleansing</i>	36
4.2.2	<i>Tokenization</i>	36
4.2.3	<i>Case Folding</i>	37
4.2.4	<i>Stemming</i>	38
4.2.5	<i>Stopword Removal</i>	38
4.3	HASIL PENELITIAN DENGAN <i>SUPPORT VECTOR MACHINE</i>	39
4.3.1	Evaluasi Model Setiap Fold	39
4.3.2	Evaluasi Model Berdasarkan Aspek	41
4.4	PEMBAHASAN.....	42
4.5	APLIKASI HASIL PENELITIAN.....	43
BAB V KESIMPULAN DAN SARAN.....		44
5.1	KESIMPULAN.....	44
5.2	SARAN	44
DAFTAR PUSTAKA		46
LAMPIRAN		49

DAFTAR TABEL

TABEL 2.1 PENELITIAN TERDAHULU	19
TABEL 3.1 JADWAL KEGIATAN PENELITIAN	23
TABEL 3.2 HARDWARE.....	24
TABEL 3.3 SOFTWARE.....	24
TABEL 3.4 CONTOH PELABELAN MANUAL	26
TABEL 3.5 CLEANSING.....	27
TABEL 3.6 TOKENIZATION.....	28
TABEL 3.7 CASE FOLDING	28
TABEL 3.8 STEMMING.....	29
TABEL 3.9 STOPWORD REMOVAL.....	30
TABEL 3.10 CONTOH K-FOLD CROSS VALIDATION.....	31
TABEL 4.1 PERBANDINGAN SETELAH DI CLEANSING	36
TABEL 4.2 PERBANDINGAN DATA CLEANSING DENGAN TOKENIZATION	37
TABEL 4.3 PERBANDINGAN DATA TOKENIZATION DAN CASE FOLDING.....	37
TABEL 4.4 PERBANDINGAN DATA CASE FOLDING DAN STEMMING	38
TABEL 4.5 PERBANDINGAN DATA STEMMING DAN STOPWORD REMOVAL	39
TABEL 4. 6 HASIL EVALUASI MODEL SETIAP FOLD	39
TABEL 4. 7 MATRIKS HASIL EVALUASI PER ASPEK	41

DAFTAR GAMBAR

GAMBAR 1.1	PODIUM YANG HANYA DENGAN LAMPU MOBIL	1
GAMBAR 1.2	KAMAR TIDUR ATLET INDONESIA	2
GAMBAR 2.1	KERANGKA BERPIKIR.....	22
GAMBAR 3.1	FLOWCHART PENELITIAN.....	25
GAMBAR 4.1	HASIL CRAWLING DATA TWITTER DENGAN TWEETPY.....	33
GAMBAR 4.2	HASIL DARI FILTERING.....	34
GAMBAR 4.3	DATA SETELAH PELABELAN MANUAL	34
GAMBAR 4.4	GAMBAR PERBANDINGAN DATA POSITIF DAN NEGATIF	35
GAMBAR 4.5	PERBANDINGAN JUMLAH ASPEK	35
GAMBAR 4.6	GRAFIK HASIL EVALUASI MODE FOLD	41
GAMBAR 4.7	GRAFIK EVALUASI PER ASPEK.....	42



DAFTAR LAMPIRAN

LAMPIRAN 1. LEMBAR PERSETUJUAN DOSEN	49
LAMPIRAN 2. SURAT PERNYATAAN DOSEN PEMBIMBING 1.....	49
LAMPIRAN 3. SURAT PERNYATAAN DOSEN PEMBIMBING 2.....	51
LAMPIRAN 4. LEMBAR KONSULTASI DOSEN PEMBIMBING 1	52
LAMPIRAN 5. LEMBAR KONSULTASI DOSEN PEMBIMBING 2	53
LAMPIRAN 6. SCREENSHOT KODE PEMBUATAN MODEL.....	54



BAB I

PENDAHULUAN

1.1 Latar belakang

Southeast Asian Games atau *SEA Games* merupakan festival multi-olahraga untuk Kawasan di Asia Tenggara yang dilaksanakan setiap dua tahun sekali (Utami, 2021). Acara *SEA games* diadakan pertama kali diadakan pada 12 hingga 17 Desember 1959 di Bangkok, Thailand. Hingga sekarang tahun 2023, *SEA Games* sudah diadakan hingga 32 kali. Di mana tahun ini negara Cambodia/Kamboja menjadi tuan rumah untuk *SEA Games* ke 32.

Kamboja, sebagai salah satu negara pelopor penyelenggaraan *SEA Games*, resmi menjadi tuan rumah *SEA Games* untuk pertama kalinya pada tahun 2023. Hal ini merupakan momen bersejarah bagi Kamboja, sekaligus tantangan besar untuk mempersiapkan penyelenggaraan yang berkesan bagi masyarakat Kamboja, atlet-atlet yang akan bertanding, maupun penonton dari luar negara.

Kesiapan fasilitas penyelenggaraan *SEA Games* 2023 di Kamboja menjadi sorotan publik. Hal ini disebabkan oleh beredarnya foto-foto fasilitas yang diterima atlet Indonesia yang terlihat kurang matang seperti yang ditampilkakan pada gambar 1.1 hingga 1.4 dibawah ini.



Gambar 1.1 Podium yang hanya dengan lampu mobil

Gambar tersebut diambil dari salah satu pengguna akun Twitter yang bernama @ainurohman. Kejadian tersebut terjadi pada cabang olahraga lari cepat. Dimana perwakilan Indonesia meraih juara pertama dan akan dilakukan pemberian medali,

namun sayangnya lampu podium mati dan digantikan dengan lampu mobil sebagai sumber cahaya utama. Post tersebut mengundang berbagai macam tanggapan dari warganet salah satunya datang dari warganet yang berkomentar sebagai berikut.

“Itu niat apa ngga sih jadi tuan rumah? Kok lighting nya pake lampu mobil



Komentar tersebut berasal dari akun @Ravico27. Komentar tersebut menggambar kekecewaan warganet tentang fasilitas yang disediakan pihak penyelenggara SEA games.



Gambar 1.2 Kamar Tidur atlet Indonesia

Gambar 1.2 memperlihatkan adanya kebocoran pada kamar atlet Indonesia. Hal ini terjadi karena adanya hujan lebat yang menerjang dan mengakibatkan terjadinya bocor pada kamar atlet Indonesia yang berada di lantai 3. Berikut salah satu tanggapan warganet Twitter tentang insiden yang terjadi.

"Miris lihat kondisi kamar atlet Indonesia. Ini SEA Games loh, bukan turnamen tarkam!"

Komentar tersebut berasal dari akun @ndagels yang merasa prihatin dengan kondisi kamar yang ditempati atlet Indonesia. Dia juga membandingkan acara SEA games dengan turnamen tarkam (antar kampung).

Gambar-gambar tersebut menunjukkan kondisi fasilitas ketika penyerahan medali juara di podium dan kamar atlet Indonesia yang terlihat disiapkan dengan kurang matang. Padahal, kesiapan fasilitas merupakan salah satu faktor penting

dalam penyelenggaraan sebuah acara olahraga besar. Fasilitas yang memadai dapat memberikan kenyamanan dan keamanan bagi atlet maupun penonton, sehingga dapat meningkatkan kualitas penyelenggaraan acara.

Komentar warganet tentang fasilitas SEA Games 2023 di Kamboja cukup beragam, mulai dari positif bahkan negatif. Adanya berbagai macam respon ini penting untuk di analisis sehingga pihak penyelenggara SEA Games selanjutnya dapat mengambil langkah yang tepat. Salah satu cara yang bisa digunakan untuk menganalisis data tersebut adalah dengan menggunakan metode analisis sentimen.

Analisis Sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negatif (Pang & Lee, 2008). Dengan menggunakan analisis sentimen dapat mengelompokkan polaritas yang terdapat dalam suatu teks seperti di dalam dokumen, kalimat, maupun fitur tingkat aspek yang dapat diketahui apakah pendapat tersebut bersifat positif, negatif, atau netral (Nugraha, Harani, & Habibi, 2020). Namun di dalam sebuah ulasan dapat berisi beberapa aspek yang masing-masing mengandung polaritas yang berbeda, sehingga akan membuat sebuah kalimat tidak sepenuhnya hanya mengandung satu jenis polaritas (Pamungkas & Romadhony, 2021)

Melakukan proses Analisis Sentimen perlu menggunakan teknik tertentu, salah satunya dengan menggunakan *machine learning*. *Support Vector Machine* (SVM) merupakan salah satu Algoritma yang dapat digunakan untuk pemrosesan Analisis Sentimen. SVM merupakan suatu teknik untuk menemukan *hyperplane* yang bisa memisahkan dua set data dari dua kelas yang berbeda (Ladayya, Siregar, Pranoto, & Muchtar, 2022).

Penelitian terkait pernah dilakukan oleh Ikram Maulana, Winda Apriandari, dan Agung Pambudi pada tahun 2023 dengan judul “Analisis Sentimen berbasis Aspek terhadap ulasan aplikasi MyPertamina menggunakan *support vector machine*”. Penelitian ini menggunakan ulasan Google Play Store tentang aplikasi MyPertamina sebagai objek penelitian. Penelitian tersebut menggunakan 19562 ulasan dari bulan Agustus 2022 sampai Maret 2023. Menggunakan 2 metode klasifikasi yaitu klasifikasi sentimen dan aspek. Klasifikasi sentimen menggunakan

2 kelas yaitu positif dan negatif. Sedangkan untuk yang klasifikasi aspek yang digunakan dalam penelitian tersebut terdiri dari 3 aspek yaitu Bug, Kegunaan, dan Pembayaran. Kemudian model yang digunakan adalah SVM (*Support Vector Machine*), dari penelitian yang telah dilakukan didapatkan hasil untuk klasifikasi sentimen dan klasifikasi berbasis aspek berurutan menunjukkan akurasi sebesar 92% dan 96% (Maulana, Apriandari, & Pambudi, 2023).

Penelitian serupa juga dilakukan oleh Yutika, Adiwijaya, dan Faraby yang di publikasi pada tahun 2021. Pada penelitian ini mereka membahas tentang Analisis Sentimen dari *review* forum Female Daily berbasis Aspek menggunakan metode TF-IDF dan Naive Bayes. Penelitian ini menggunakan 4 kategori aspek yaitu Harga, Kemasan, Produk, dan Aroma, dengan klasifikasi sentimen terdiri dari positif, netral, dan negatif. Dengan menggunakan dataset sebanyak 5054, dilakukan sebanyak 3 Evaluasi Skenario untuk mengetahui performansi setiap dataset ketika tidak di terjemahkan dan sesudah di terjemahkan serta pengaruh *preprocessing*, *stopword*, dan *stemming*. Pada Skenario pertama, dataset diterjemahkan dan didapatkan hasil *F1-Score* tertinggi yaitu 62,81%. Pada skenario kedua, dataset tidak menggunakan tahapan *stopword removal* dan di dapatkan *F1-Score* tertinggi sebesar 62,81%. Pada skenario ketiga, digunakan *hyperparameter tuning* didapatkan hasil *F1-Score* tertinggi sebesar 61,86%. Kemudian disimpulkan bahwa dataset yang telah di terjemahkan dengan tanpa *stopword removal* dan parameter alpha sebesar 1, *min_df* sebesar 0,01, *max_df* sebesar 0,7, dan *max_features* sebesar 2000 mendapatkan hasil terbaik yaitu *F1-Score* sebesar 62,81% (Yutika, Adiwijaya, & Faraby, 2021).

Melihat Penelitian-penelitian terkait, mendorong dilakukannya penelitian Analisis Sentimen berbasis aspek dari Pengguna Twitter mengenai fasilitas SEA Games 2023. Penelitian ini akan melakukan klasifikasi aspek yang kemudian membagi kelas menjadi kenyamanan, kelengkapan, dan kebersihan.

Tujuan penelitian ini adalah untuk mendapatkan pemahaman yang lebih mendalam tentang persepsi publik terhadap kualitas fasilitas dalam penyelenggaraan SEA Games 2023. Analisis sentimen dari pengguna Twitter dilakukan untuk mengidentifikasi tren, pola, serta *Tweet* yang muncul terkait

fasilitas tersebut. Hasil analisis ini diharapkan dapat memberikan wawasan yang berharga bagi penyelenggara dan pihak terkait dalam meningkatkan kualitas fasilitas, merespons masalah yang mungkin muncul, dan memastikan pengalaman yang memuaskan bagi atlet dan penonton..

Analisis sentimen fasilitas SEA Games 2023 dilakukan dengan menggunakan data dari API platform Twitter sebagai sumber informasi yang relevan. Python dimanfaatkan sebagai alat pengolahan data untuk menganalisis sentimen dan aspek yang terdapat pada data *Tweet* pengguna Twitter secara efisien. Hasil analisis memberikan gambaran yang komprehensif tentang pandangan dan perasaan terkait dengan fasilitas SEA Games 2023.

1.2 Identifikasi masalah

Dari latar belakang di atas, dapat dilihat bahwa masalah yang terjadi :

1. Pemberian fasilitas yang diberikan panitia dari negara Kamboja dinilai tidak siap.
2. Tersebar foto-foto yang menunjukkan ketidaksiapan fasilitas *SEA Games 2023*
3. Adanya beragam sentiment di sosial media mengenai fasilitas yang diberikan pihak penyelenggara *SEA Games 2023*.

1.3 Pembatasan masalah

Penelitian ini memiliki batasan dalam melakukan penelitian, sehingga penelitian yang lebih terfokus dan tepat sasaran. Berikut merupakan batasan yang diberikan:

1. Fokus pada analisis sentimen fasilitas *SEA Games 2023*.
2. Pengumpulan data hanya diambil melalui platform Twitter.
3. Analisis sentimen hanya terdiri dari kelas positif dan negatif.
4. Analisis aspek hanya terdiri dari kenyamanan, kelengkapan, dan kebersihan.
5. Data yang digunakan hanya *Tweet* warganet yang menggunakan bahasa Indonesia.
6. Data yang diambil memiliki rentang dari 25 April 2023 hingga 18 Mei 2023.

7. Algoritma yang digunakan adalah *Support Vector Machine*.

1.4 Rumusan Masalah

Berdasarkan latar belakang yang telah diberikan, rumusan masalah dari penelitian ini adalah bagaimana hasil Analisis Sentimen Berbasis Aspek Terhadap Fasilitas Sea Games 2023 Melalui Media Twitter Menggunakan *Algoritma Support Vector Machine*?

1.5 Tujuan penelitian

Dari latar belakang yang telah disebutkan, terdapat tujuan penelitian yaitu Untuk mengetahui hasil sentimen publik berdasarkan aspek mengenai fasilitas SEA Games 2023

1.6 Manfaat penelitian

Dari penjelasan latar belakang dan tujuan penelitian yang diberikan, diharapkan penelitian ini memberikan manfaat sebagai berikut:

1. Peningkatan kualitas fasilitas
2. Peningkatan citra dan reputasi tentang acara SEA Games
3. Pengambilan keputusan yang lebih baik
4. Penyempurnaan penyelenggaraan SEA Games berikutnya
5. Data yang digunakan hanya *Tweet* yang menggunakan bahasa Indonesia.

BAB II

KERANGKA TEORITIK

2.1 Kerangka Teoritik

2.1.1 SEA Games

Southeast Asian Games atau lebih dikenal sebagai *SEA Games*, merupakan sebuah festival multi-olahraga yang di khususkan untuk kawasan Asia Tenggara yang diadakan setiap dua tahun sekali. Hingga saat ini acara *SEA Games* sudah diberlangsungkan sebanyak 32 kali.

SEA Games pertama kali digelar tahun 1959 di Thailand dengan jumlah peserta sebanyak 6 negara. Indonesia sendiri baru bergabung pada tahun 1977 bersama dengan negara Filipina. Hingga saat ini sudah ada 11 negara yang menjadi peserta dalam acara *SEA Games* tersebut yaitu Republik Indonesia, Brunei Darussalam, Kerajaan Kamboja, Republik Demokratik Laos, Malaysia, Uni Myanmar, Republik Filipina, Republik Singapura, Kerajaan Thailand, dan Republik Sosialis Vietnam (Kompas, 2021).

2.1.2 Twitter

Twitter didirikan pada tahun 2006 oleh Jack Dorsey, Noah Glass, Biz Stone, dan Evan Williams. Platform ini awalnya dirancang sebagai layanan mikroblogging, yang memungkinkan pengguna untuk mengirim pesan singkat yang disebut "tweet". Setiap tweet mampu menampung hingga 140 karakter, kemudian ditingkatkan menjadi 280 karakter pada tahun 2017.

Pada tahun 2022, Twitter diakuisisi oleh Elon Musk, seorang pengusaha miliarder. Musk mengatakan bahwa dia ingin membuat Twitter menjadi platform yang lebih bebas dan terbuka. Pada bulan oktober di tahun yang sama, musk mengumumkan bahwa akan ada pergantian nama pada Twitter menjadi X. kemudian rencana tersebut dilaksanakan dan secara resmi Twitter berganti nama menjadi X pada 31 Juli 2023 dan memiliki pengguna aktif sebesar 225 Juta pengguna pada bulan November 2023.

Melihat banyaknya pengguna internet yang menggunakan Twitter untuk berbagi informasi dan berdiskusi, membuat beberapa perusahaan memanfaatkan hal tersebut untuk mengetahui pola, dan tren yang sedang terjadi diantara pengguna X. caranya adalah dengan melakukan beberapa macam prosedur seperti berikut:

1. API Twitter

API Twitter adalah kumpulan alat dan dokumentasi yang disediakan oleh pihak resmi Twitter untuk memungkinkan *Developer* untuk mengakses dan mengontrol Twitter. Ada berbagai fungsi-fungsi yang dapat digunakan *developer* untuk mengakses data Twitter, seperti mendapatkan *Tweet*, Nama pengguna, *Hashtag*, *Retweet*, *Like*, *Reply*, *Mention* .

2. Layanan pihak ketiga

Merupakan cara yang disediakan oleh pihak ketiga yang biasanya memiliki tampilan yang mudah digunakan dari pada API Twitter. Pengguna data mengunduh data Twitter dalam bentuk seperti CSV dan JSON sehingga mudah digunakan. Berikut merupakan beberapa layanan pihak ketiga yang umum digunakan dalam mengambil data Twitter Tweepy, Sprout Social, Mentionlytics, dan masih banyak lagi.

3. Alat Scraping

Web Scraping adalah metode untuk mengekstraksi data dari *Worldwide Web* (WWW) dan menyimpannya ke sistem file atau *database* untuk pengambilan atau analisis (Hafiz & Sudarmilah, 2023). Sedangkan *Alat Scraping* adalah sebuah *software* yang membantu menjalankan proses *Scraping*. Contoh alat *scraping* BeautifulSoup. Tweet Harvest, Scrapy.

2.1.3 Analisis Sentimen

Analisis Sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negatif (Pang & Lee, 2008). Analisis Sentimen dilakukan dengan tujuan menganalisa pendapat, sentimen, evaluasi, sikap, penilaian dan emosi seseorang. apakah pembicara atau penulis berkenan dengan suatu topik, produk, layanan, organisasi, individu, atau kegiatan tertentu (Ipmawati, Kusriani, & Luthfi, 2017).

2.1.4 Analisis Sentimen Berbasis Aspek

Analisis Sentimen Berbasis Aspek atau *Aspect-Based Sentiment Analysis* (ABSA) merupakan salah satu bagian dari *opinion mining* yang bertujuan untuk mendeteksi polaritas teks tertulis berdasarkan aspek tertentu (Rahmayanti, 2022). Analisis Sentimen Berbasis Aspek memungkinkan perusahaan untuk memperoleh pemahaman yang lebih dalam tentang produk dan layanan tertentu dengan cepat, mudah, dan fokus pada kebutuhan dan harapan pelanggan mereka. (Giffari, 2022)

Di dalam metode ABSA, aspek dapat dibagi menjadi dua tingkat yaitu istilah aspek (juga disebut target aspek) dan kategori aspek (Liu, Shen, Zhenjiang, Zhang, & Mi, 2019). Istilah aspek (Target aspek) merupakan kata dalam suatu teks/opini yang mengacu pada entitas yang lebih spesifik yang sedang di bahas. Sedangkan target aspek merupakan sebuah pengelompokan dari target aspek yang lebih luas. Kategori aspek ini mencerminkan karakteristik entitas yang sedang dibahas secara umum. Contoh kalimat berikut:

“Fasilitas SEA *Games* 2023 di Kamboja cukup bagus, terutama dari segi akomodasi dan transportasi. Namun, masih ada beberapa hal yang perlu ditingkatkan, seperti fasilitas kesehatan dan keamanan.”

Berdasarkan contoh kalimat diatas kata “akomodasi”, “transportasi”, “kesehatan” dan “keamanan” termasuk kedalam istilah aspek. Sedangkan kategori aspek-nya adalah “fasilitas”. Dapat terlihat juga di dalam 1 kalimat tersebut terdapat 2 sentimen yaitu sentimen positif untuk istilah aspek “akomodasi” dan “transportasi”, dan sentimen negatif untuk istilah aspek “kesehatan” dan “keamanan”.

2.1.5 *Preprocessing*

Preprocessing merupakan tahap yang berfungsi untuk proses pembersihan dan persiapan teks untuk klasifikasi (Emma, Liu, & Shi, 2013). Proses dari *Preprocessing* sering dilakukan sebelum pembuatan model *machine learning*, dengan adanya proses ini diharapkan waktu, keakuratan, dan kompleksitas data yang digunakan menjadi lebih optimal. Berikut merupakan beberapa metode *preprocessing* yang akan digunakan dalam penelitian ini.

1. *Cleansing*

Merupakan proses dimana *dataset* akan dilakukan pembersihan *emoticon*, angka, *whitespace*, dan tanda baca (Giffari, 2022). Hal ini dilakukan untuk menghilangkan *noise* sehingga informasi yang diberikan akan lebih akurat dan mudah diproses. Contoh:

“Film ini sangat bagus! Aktingnya luar biasa, ceritanya menarik, dan efek visualnya menakjubkan. Saya sangat merekomendasikannya!”

Setelah dilakukan *Cleansing*

“Film ini sangat bagus Aktingnya luar biasa ceritanya menarik dan efek visualnya menakjubkan Sata sangat merekomendasikannya”

2. *Tokenization*

Tokenization adalah proses memecah teks menjadi bagian yang lebih kecil, yang disebut token. Token dapat berupa kata, frasa, atau bahkan kalimat individual (Bird, Klein, & Loper, 2009). *Tokenization* kata akan merubah kalimat menjadi bagian-bagian kata untuk dijadikan token Contoh:

“Film ini sangat bagus!”

Setelah dilakukan *Tokenization*

“[Film], [ini], [sangat], [bagus], [!]”

Tokenization frasa akan merubah kalimat menjadi beberapa frasa berdasarkan maknanya untuk dijadikan token, contoh:

"Dia membeli buku baru di toko buku."

Setelah dilakukan Tokenization frasa akan menjadi

“[Dia membeli], [buku baru], [di toko baru]”

Kemudian tokenization kalimat individual akan merubah kalimat majemuk menjadi beberapa bagian kalimat individual untuk dijadikan token, contoh:

“Budi sedang belajar bahasa Inggris. Dia sangat antusias dan ingin menguasainya. Bahasa Inggris merupakan bahasa yang penting di era globalisasi.”

Setelah dilakukan Tokenization kalimat akan menjadi

“[Budi sedang belajar bahasa inggris.], [Dia sangat antusias dan ingin menguasainya.], [Bahasa Inggris merupakan bahasa yang penting di era globalisasi.]”

Setelah melihat beberapa contoh dari penggunaan tokenization, dipilihlah tokenization kata sebagai metode tokenization yang akan digunakan dalam penelitian ini.

3. *Case Folding*

Case Folding merupakan tahapan dimana seluruh teks pada *dataset* diubah menjadi huruf kecil (*lowercase*) untuk mempermudah proses pengolahan (Subowo, Artanto, Putri, & Umaedi, 2022). Dengan menggunakan *Case Folding* semua kata-kata yang mengandung huruf kapital akan diubah menjadi huruf kecil untuk mempermudah proses analisis. Contoh:

“Kertas HVS itu mahal sekali”

Setelah dilakukan *Case Folding*

“kertas hvs itu mahal sekali”

4. *Stemming*

Stemming adalah proses untuk mengurangi semua kata dengan batang kata yang sama menjadi bentuk umum. Tujuan utama *stemming* adalah untuk meningkatkan efisiensi pencarian informasi dengan mengurangi jumlah kata yang harus diindeks dan dicari (Balakrishnan & Yemoh, 2014). Contoh:

“Saya sedang membaca buku tentang sejarah Indonesia.”

Setelah dilakukan *Stemming*

“saya sedang baca buku tentang sejarah indonesia.”

5. *Stopword Removal*

Stopword Removal yaitu penghapusan kata-kata yang dianggap umum dan tidak memberikan banyak informasi dalam analisis teks (Arum, Butsianto, & Astuti, 2023). Kata-kata yang dimaksud seperti kata “dan”, “di”, “atau”, “ada”, dan lain-lain. Tujuan dari proses ini adalah untuk mengurangi jumlah kata yang disimpan dalam daftar token yang nantinya akan dilakukan proses selanjutnya (Septian, Fahrudin, & Nugroho, 2019). Contoh:

“Saya ingin membeli sepatu baru di toko online terpercaya dengan harga murah.”

Setelah dilakukan *Stopword Removal*

“ingin membeli sepatu baru toko online terpercaya harga murah”

2.1.6 Support Vector Machine

Support Vector Machine (SVM) memiliki prinsip dasar *linear classifier*, yang berarti SVM dapat bekerja pada kasus klasifikasi yang secara linier dapat dipisahkan. Namun, SVM telah dikembangkan agar dapat bekerja pada problem non-linier dengan memasukkan konsep *kernel* pada ruang kerja berdimensi tinggi. (Octaviani, Wilandari, & Ispriyanti, 2014). Pada ruang berdimensi tinggi, *Support Vector Machine* (SVM) merupakan suatu teknik untuk menemukan *hyperplane* yang bisa memisahkan dua set data dari dua kelas yang berbeda (Ladayya, Siregar, Pranoto, & Muchtar, 2022).

hyperplane terbaik adalah dengan mencari *hyperplane* yang terletak di tengah-tengah antara dua bidang pembatas kelas dan untuk mendapatkan *hyperplane* terbaik itu, sama dengan memaksimalkan *margin* atau jarak antara dua set objek dari kelas yang berbeda.

SVM yang terjadi pada klasifikasi linier memiliki *hyperplane* dengan notasi:

$$f(x) = W^T x + b \quad (1)$$

Sehingga akan diperoleh persamaan:

$$[W^T \cdot x_i + b] \geq 1 \text{ untuk } y_i = +1 \quad (2)$$

$$[W^T \cdot x_i + b] \leq -1 \text{ untuk } y_i = -1 \quad (3)$$

Keterangan:

W : bobot *hyperplane*

x_i : himpunan data training ke- i

b : bias *hyperplane*

y_i : label kelas dari data x_i

Untuk mendapatkan *hyperplane* terbaik perlu di lakukan suatu metode yang dinamakan metode *Quadratic Programming* (QP) yaitu dengan meminimalkan $W^T W$ dengan fungsi batasan $y_i(W^T \cdot x_i + b) \geq 1$. Masalah tersebut dapat diselesaikan menurut (Ladayya, Siregar, Pranoto, & Muchtar, 2022) dengan menggunakan formula langrange berikut ini:

$$L_p(W, b, \alpha) = \frac{1}{2} W^T W - \sum_{i=1}^n \alpha_i [y_i(x_i^T W + b) - 1] \quad (4)$$

Keterangan:

W : bobot *hyperplane*

x_i : himpunan data training ke- i

b : bias *hyperplane*

y_i : label kelas dari data x_i

α : lagrange multiplier

Dengan meminimumkan $L_p(W, b, \alpha)$ terhadap W dan b maka dapat mengoptimalkan nilai dari persamaan (3) sehingga lagrangian untuk dual problem sebagai berikut:

$$L_D(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (5)$$

Data *training* dengan $\alpha_i > 0$ terletak pada *hyperplane* disebut *support vector*. Data *training* yang tidak berada pada *hyperplane* tersebut mempunyai $\alpha_i = 0$. Kemudian menurut (Octaviani, Wilandari, & Ispriyanti, 2014) ditentukan kelas dari data training dengan fungsi berikut:

$$f(x_t) = \sum_{s=1}^{ns} \alpha_s y_s x_s \cdot x_t + b \quad (6)$$

Keterangan:

x_t : data testing

x_s : data support vector

ns : banyak data support vector

Selain SVM yang bekerja pada data linier, dengan bantuan konsep *kernel*. SVM dapat menyelesaikan masalah klasifikasi secara *non-linear*. Berikut beberapa fungsi yang dapat membentuk matriks *kernel* yang umum di gunakan dalam menggunakan SVM:

1. Kernel Polinomial:

$$K(x, y) = (x^T y + 1)^d \quad (7)$$

2. Kernel RBF:

$$K(x, y) = \exp(-\gamma \|x - y\|^2), \gamma > 0 \quad (8)$$

3. Kernel Sigmoid:

$$K(x, y) = \tanh(\gamma(x^T y) + r) \quad (9)$$

2.1.7 K-Fold Cross Validation

Cross validation adalah sebuah teknik validasi model untuk menilai bagaimana hasil statistik analisis akan menggeneralisasi kumpulan data independen. Teknik ini utamanya digunakan untuk melakukan prediksi model dan memperkirakan seberapa akurat sebuah model prediktif ketika dijalankan dalam praktiknya (Azis, Purnawansyah, Fattah, & Putri, 2020). Sedangkan *K-Fold Cross Validation* adalah cabang yang lebih spesifik dari *Cross Validation* yang berfungsi untuk menilai kinerja proses sebuah algoritma serta digunakan untuk mengurangi waktu dalam proses klasifikasi dengan tetap menjaga nilai keakuratan dengan membagi data secara acak dan mengelompokkan data sebanyak *K-Folds* agar tidak terjadi *overlapping* pada data *testing* (Apriyani & Kurniati, 2020). Kemudian, dicarikan nilai akurasi rata-rata untuk melihat seberapa akurat model memprediksi data dengan rumus:

$$\text{akurasi rata - rata} = \frac{\sum \text{akurasi } k}{k} \quad (10)$$

Keterangan:

k: jumlah fold

2.1.8 Google Collab

Colaboratory, atau singkatnya disebut “Colab”, merupakan produk dari Google Research. Colab memungkinkan siapa saja menulis dan mengeksekusi kode python arbitrer melalui *browser*, dan sangat cocok untuk *machine learning*, analisis data, serta pendidikan. Secara lebih teknis, Colab merupakan layanan Jupyter notebook yang di-*hosting* dan dapat digunakan tanpa instalasi, serta menyediakan akses tanpa biaya ke *resource* komputasi termasuk GPU (Google, 2017).

2.1.9 Tweepy

Tweepy adalah library open-source yang dibuat dengan Python untuk berinteraksi dengan Twitter API. *Software* ini dikembangkan oleh Joshua Roesslein bersama komunitas *developer open-source* lainnya. Tweepy memiliki berbagai fungsi seperti mengambil data Twitter, membuat post, hingga mengikuti akun lainnya.

Dikarenakan berbagai macam fungsinya, Tweepy menjadi salah satu *library open-source* python yang banyak digunakan di *platform* Twitter. Penggunaan Tweepy bisa untuk berbagai macam keperluan seperti analisis sentimen, pembuatan bot, analisis jaringan, dan lain sebagainya.

2.2 Penelitian Terdahulu

Dalam Penelitian, penting diperlukan sebuah sumber penelitian terdahulu sebagai acuan. Sumber penelitian-penelitian ini berfungsi sebagai untuk meminimalisir terjadinya plagiarisme dan sebagai arahan dalam melakukan penelitian. Analisis Sentimen Berbasis Aspek merupakan salah satu pendalaman dari Analisis Sentimen yang mana berfokus pada sentimen setiap aspek yang diteliti, sehingga topik ini menjadi semakin banyak dilakukan karena lebih terfokus daripada Analisis Sentimen pada umumnya.

Penelitian dengan tema Analisis Sentimen Berbasis Aspek juga pernah dilakukan oleh Jessica Widyadhana Iskandar, dan Yessica Nataliani yang dilakukan pada tahun 2021 yang berjudul “Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek”. Penelitian ini bertujuan untuk mengetahui sentimen publik tentang salah satu *gadget* yaitu Samsung Galaxy Z Flip

3 melalui media komentar Youtube. Setelah dilakukan pengambilan data melalui komentar, didapatkan data sebesar 9.597 komentar yang kemudian diolah kembali dengan model *Cross-Industry Standard Process for Data Mining* (CRISP-DM) dan di dapatkan hasil 1.391 komentar yang relevan dengan topik penelitian. Data yang sudah di modelkan kemudian data dikelompok-kan berdasarkan aspek desain, harga, aspek, dan citra merk. Menggunakan metode klasifikasi *Naïve bayes* (NB), *Support Vector Machine* (SVM), dan *k-Nearest Neighbor* (k-NN) data tersebut diolah dan dibandingkan. Didapatkan hasil bahwa dari ketiga metode klasifikasi tersebut, SVM merupakan metode yang terbukti mendapatkan hasil yang terbaik dilihat dari keempat aspek. Rata-rata *accuracy* SVM sebesar 96,43%, yaitu aspek desain sebesar 94.40%, aspek harga sebesar 97.44%, aspek spesifikasi sebesar 96.22%, dan aspek citra merk sebesar 97.63% .

Penelitian kedua dilakukan oleh Ikram Maulana, Winda Apriandari, dan Agung Pambudi pada tahun 2023 dengan judul “Analisis Sentimen berbasis Aspek terhadap ulasan aplikasi MyPertamina menggunakan *support vector machine*”. Penelitian ini menggunakan ulasan Google Play Store tentang aplikasi MyPertamina sebagai objek penelitian. Penelitian tersebut menggunakan 19562 ulasan dari bulan Agustus 2022 sampai Maret 2023. Menggunakan 2 metode klasifikasi yaitu klasifikasi sentimen dan aspek. Klasifikasi sentimen menggunakan 2 kelas yaitu positif dan negatif. Sedangkan untuk yang klasifikasi aspek yang digunakan dalam penelitian tersebut terdiri dari 3 aspek yaitu Bug, Kegunaan, dan Pembayaran. Kemudian model yang digunakan adalah SVM (*Support Vector Machine*), dari penelitian yang telah dilakukan didapatkan hasil untuk klasifikasi sentimen dan klasifikasi berbasis aspek berurutan menunjukkan akurasi sebesar 92% dan 96%.

Penelitian ketiga juga dilakukan oleh yutika, Adiwijaya, dan Faraby yang di publikasi pada tahun 2021. Pada penelitian ini mereka membahas tentang “Analisis Sentimen dari *review* forum Female Daily berbasis Aspek menggunakan metode TF-IDF dan *Naive Bayes*”. Penelitian ini menggunakan 4 kategori aspek yaitu Harga, Kemasan, Produk, dan Aroma, dengan klasifikasi sentimen terdiri dari positif, netral, dan negatif. Dengan menggunakan *dataset* sebanyak 5054, dilakukan sebanyak 3 Evaluasi Skenario untuk mengetahui performansi setiap dataset ketika

tidak di terjemahkan dan sesudah di terjemahkan serta pengaruh *preprocessing*, *stopword* dan *stemming*. Pada Skenario pertama, dataset diterjemahkan dan didapatkan hasil F1-Score tertinggi yaitu 62.81%. Pada skenario kedua, *dataset* tidak menggunakan tahapan *stopword removal* dan di dapatkan F1-Score tertinggi sebesar 62.81%. Pada skenario ketiga, digunakan *hyperparameter tuning* didapatkan hasil *F1-Score* tertinggi sebesar 61.86%. Kemudian disimpulkan bahwa *dataset* yang telah di terjemahkan dengan tanpa *stopword removal* dan parameter *alpha* sebesar 1, *min_df* sebesar 0.01, *max_df* sebesar 0.7, dan *max_features* sebesar 2000 mendapatkan hasil terbaik yaitu *F1-Score* sebesar 62.81%

Penelitian oleh Rima Tamara Aldisa, dan Pandu Maulana pada 2022 dengan judul “Analisis Sentimen Opini Masyarakat terhadap Vaksinasi Booster COVID-19 dengan perbandingan metode Naïve Bayes, Decision Tree, dan SVM”. Dengan menggunakan 10500 data opini dari Twitter sepanjang tahun 2021, menghasilkan perhitungan untuk setiap dataset positif bernilai 2248, dan untuk negatif bernilai 751. Kemudian dengan metode *Hold Out* didapatkan SVM memiliki tingkat akurasi, *recall*, F1, dan AUC terbesar yaitu 83.3%, 99.8%, 88.8%, dan 75.4%. Sedangkan *Precision* terbaik di dapatkan oleh model Naïve Bayes berjumlah 83.8%.

Penelitian terakhir yang di jadikan sebagai sumber penelitian terdahulu berasal dari jurnal karya Mar Atul A. T. U, Pika Silvianti, dan Mohammad Masjkur dengan judul “Algoritma *Support Vector Machine* untuk Analisis Sentimen Berbasis Aspek Ulasan Game Online Mobile Legends: Bang-Bang” yang diterbitkan pada tahun 2023. Penelitian ini menggunakan 2295 ulasan yang berasal dari Ulasan Playstore terhadap aplikasi Mobile Legends: Bang-Bang. Dengan mengklasifikasi sentimen positif dengan lebel (1), negatif dengan lebel (-1), dan netral dengan lebel (0), kemudian diklasifikasikan juga berdasarkan aspek Gameplay, Performa, Visualisasi, dan *Player*. Berdasarkan Evaluasi Model dengan *Confusion Matrix* didapatkan Visualisasi menjadi aspek dengan F1-Score positif tertinggi dengan nilai 92%. Sedangkan *Player* merupakan aspek dengan F1-Score negatif, F1-Score rata-rata, dan akurasi yang tertinggi dengan nilai berturut-turut sebagai berikut: 97%, 92%, dan 96%.

Tabel 2.1 Penelitian Terdahulu

Nama Peneliti	Jurnal	Metode Penelitian	Hasil	Judul Penelitian
Jessica Widyadhana Iskandar, dan Yessica Nataliani	Jurnal RESTI Vol. 5 No. 6 (2021) 1120 - 1126	<i>Naïve bayes (NB), Support Vector Machine (SVM), dan k-Nearest Neighbor (k-NN)</i>	SVM merupakan metode yang terbukti mendapatkan hasil yang terbaik dilihat dari keempat aspek. Rata-rata <i>accuracy</i> SVM sebesar 96.43%, yaitu spek desain sebesar 94.40%, aspek harga sebesar 97.44%, aspek spesifikasi sebesar 96.22%, dan aspek citra merk sebesar 97.63%.	Perbandingan <i>Naïve Bayes</i> , SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek
Ikram Maulana, Winda Apriandari, dan Agung Pambudi	Jurnal IDEALIS Vol. 6 NO. 2 (2023) 172 - 181	<i>Support Vector Machine (SVM)</i>	penelitian yang telah dilakukan didapatkan hasil untuk klasifikasi sentimen dan klasifikasi berbasis aspek berurutan menunjukkan akurasi sebesar 92% dan 96%.	Analisis Sentimen berbasis Aspek terhadap ulasan aplikasi MyPertamina menggunakan support vector machine
Clarisa Hasya Yutika. Adiwijaya, Said Al Faraby	Jurnal Media Informatika Budidarma Vol. 5 No. 2 (2021) 422 - 430	<i>Naïve Bayes (NB)</i>	dataset yang telah di terjemahkan dengan tanpa <i>stopword removal</i> dan parameter alpha sebesar 1, min_df sebesar 0.01, max_df sebesar 0.7, dan max_features sebesar 2000 mendapatkan hasil terbaik yaitu <i>F1-Score</i> sebesar 62.81%	Analisis Sentimen dari <i>review</i> forum Female Daily berbasis Aspek menggunakan metode TF-IDF dan <i>Naive Bayes</i>
Rima Tamara Aldisa, dan Pandu Maulana	Jurnal BITS Vol. 4 No. 1 (2022) 106 - 109	<i>Naïve Bayes (NB), Decision Tree, dan Support Vector Machine (SVM)</i>	Dengan metode <i>Hold Out</i> didapatkan SVM memiliki tingkat akurasi, <i>recall</i> , F1, dan AUC terbesar yaitu 83.3%, 99.8%, 88.8%, dan 75.4%. Sedangkan <i>Precision</i> terbaik di dapatkan oleh model <i>Naïve Bayes</i> berjumlah 83.8%.	Analisis Sentimen Opini Masyarakat terhadap Vaksinasi Booster COVID-19 dengan perbandingan metode <i>Naïve Bayes</i> , <i>Decision Tree</i> , dan SVM

Mar Atul A. T. U, Pika Silvianti, dan Mohammad Masjkur	jurnal Xplore: Journal of Statistics Vol. 12 No. 1 (2023) 63 - 77	<i>Support Vector Machine (SVM)</i>	Berdasarkan Evaluasi Model dengan <i>Confusion Matrix</i> didapatkan Visualisasi menjadi aspek dengan F1-Score positif tertinggi dengan nilai 92%. Sedangkan <i>Player</i> merupakan aspek dengan F1-Score negatif, F1-Score rata-rata, dan akurasi yang tertinggi dengan nilai berturut-turut sebagai adalah 97%, 92%, dan 96%.	Algoritme Support Vector Machine untuk Analisis Sentimen Berbasis Aspek Ulasan Game Online Mobile Legends: Bang-Bang
--	---	-------------------------------------	--	--

2.3 Kerangka Berpikir

SEA Games 2023 di Kamboja, yang berlangsung dari 5 hingga 17 Mei 2023, telah menarik perhatian banyak orang di Asia Tenggara. Ajang olahraga ini diikuti oleh 11 negara di kawasan Asia Tenggara, dan mempertandingkan 37 cabang olahraga dengan 567 nomor pertandingan.

Kontingen Indonesia, yang merupakan salah satu kontingen terbesar, berhasil meraih peringkat ketiga klasemen akhir dengan perolehan 87 medali emas, 80 medali perak, dan 109 medali perunggu. Prestasi ini merupakan pencapaian terbaik Indonesia dalam enam edisi SEA Games terakhir. Namun, di balik gemerlap prestasi para atlet Indonesia, terdapat beberapa permasalahan terkait fasilitas yang diterima atlet maupun pengunjung asal Indonesia di Kamboja. Dari permasalahan fasilitas tersebut berbagai opini publik khususnya para pengguna Twitter bermunculan. Opini-opini ini perlu dianalisa untuk mengetahui aspek-aspek apa saja yang diperlu di perbaiki dan menjadi bahan evaluasi untuk panitia SEA Games selanjutnya.

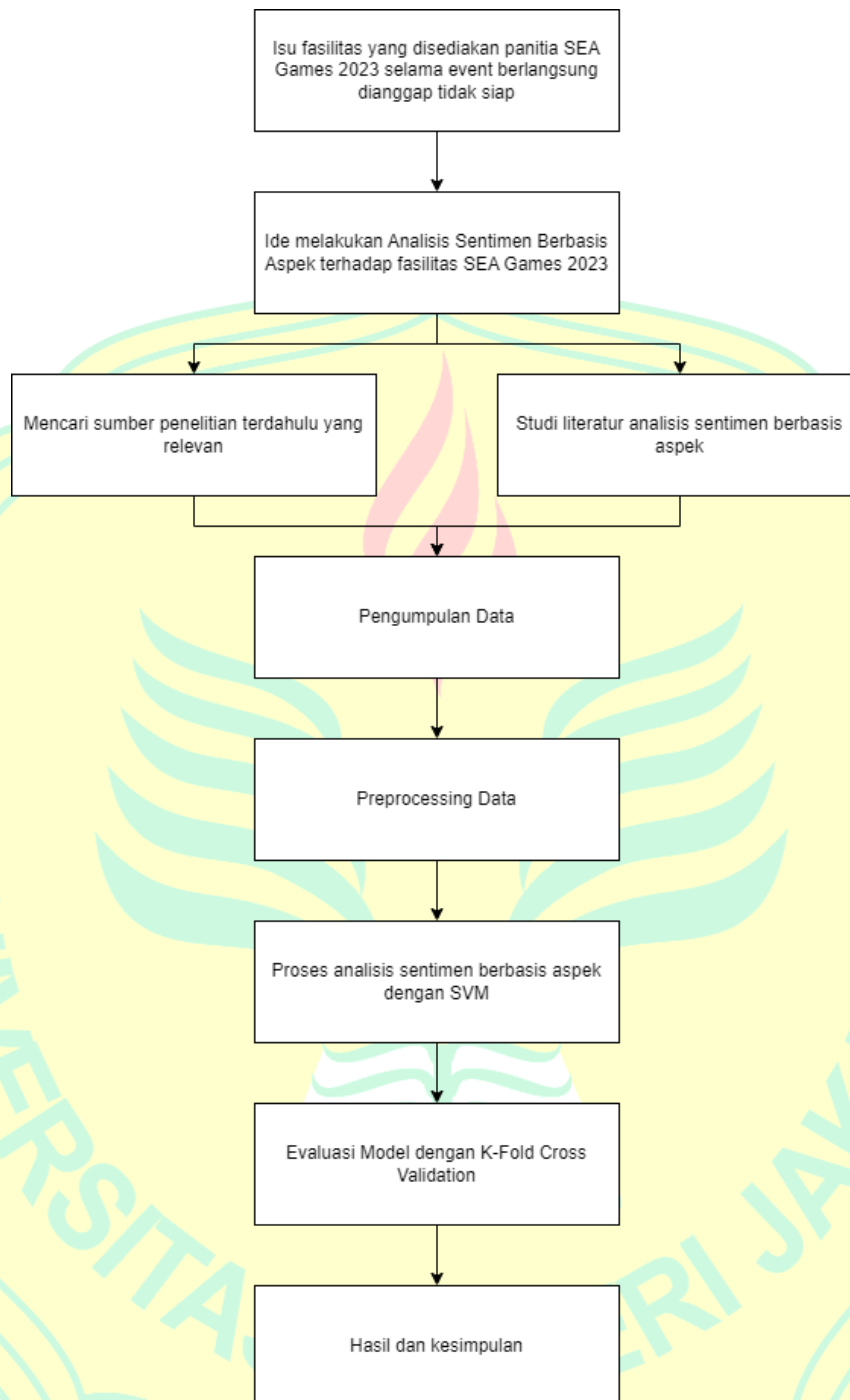
Analisis Sentimen merupakan salah satu cara untuk mengetahui bagaimana sentimen publik menanggapi fasilitas yang diberikan baik itu positif maupun negatif. Namun menggunakan metode Analisis Sentimen saja tidak cukup sehingga perlu adanya pemahan lebih dalam sehingga muncul-lah pendalaman Analisis Sentimen berbasis aspek. Yang dimana hasil dari analisis tersebut akan mengetahui aspek-aspek apa saja yang perlu diperhatikan dan ditingkatkan selanjutnya, contohnya seperti kenyamanan, kelengkapan, dan kebersihan.

Pemahaman melalui studi literatur menjadi hal yang penting dilakukan untuk mendapatkan gambaran dan pemahaman tentang hal-hal yang perlu diteliti. Teori-

teori yang perlu di cari untuk menjalankan penelitian ini dilakukan seperti pemahaman tentang *Sea Games*, *Twitter*, *Alat Scrapping*, *Analisis Sentimen*, *Analisis Sentimen Berbasis Aspek*, *Reprocessing*, *Algoritma Machine Learning*, dan *Metode Evaluasi Model*.

Setelah melakukan pemahaman melalui studi literatur, langkah berikutnya adalah mencari sumber penelitian yang relevan untuk mengetahui bagaimana proses *Analisis Sentimen Berbasis Aspek* berlangsung, *Algoritma* yang cocok untuk di gunakan untuk proses analisis data, dan metode yang cocok untuk mengevaluasi model yang sudah disiapkan. Hasil yang didapatkan setelah mencari sumber penelitian yang relevan yaitu *Support Vector Machine* menjadi salah satu *Algoritma Machine Learning* yang memiliki hasil terbaik diantara penggunaan algoritma *machine learning* yang lain seperti *Naïve Bayes*, maupun *k-Nearest Neighbor*. Kemudian *K-Fold Cross Validation* sebagai metode evaluasi model yang sudah dibuat.





Gambar 2.1 Kerangka Berpikir

BAB III METODE PENELITIAN

3.1 Tempat dan Waktu

Penelitian ini akan dilakukan secara daring dari rumah menggunakan sumber data dari platform Twitter (X) yang membahas SEA *Games* 2023. Penelitian akan dilakukan mulai Desember 2023 hingga Juni 2024. Rincian Jadwal Kegiatan dapat dilihat melalui Tabel 3.1.

Tabel 3.1 Jadwal Kegiatan Penelitian

No	Kegiatan	Waktu Penelitian						
		2023		2024				
		Des	Jan	Feb	Mar	Apr	Mei	Juni
1	Studi Literatur							
2	Penyusunan Proposal							
3	Perbaikan Proposal							
4	Pengumpulan Data							
5	Pelabelan Data							
6	Pre-processing Data							
7	Klasifikasi SVM							
8	Evaluasi Model							
9	Penyusunan Skripsi							

3.2 Alat dan Bahan Penelitian

3.2.1 Alat

Alat yang akan digunakan dalam penelitian ini adalah sebuah Laptop Lenovo Ideapad 330 dengan rincian spesifikasi yang akan di perlihatkan dalam Tabel 3.2 untuk Perangkat Keras (*Hardware*) dan Tabel 3.3 untuk Perangkat Lunak (*Software*).

Tabel 3.2 Hardware

Hardware	Spesifikasi
Prosesor	AMD A9-9425 (2.9 GHz base clock, up to 3.7 GHz max boost clock, 2 MB L2 cache)
RAM	8 GB DDR4 (Upgrade)
Penyimpanan	1 TB HDD & 125 SSD (Upgrade)
Layar	14" HD (1366x768) TN
Grafis	AMD Radeon™ R5 Graphics

Tabel 3.3 Software

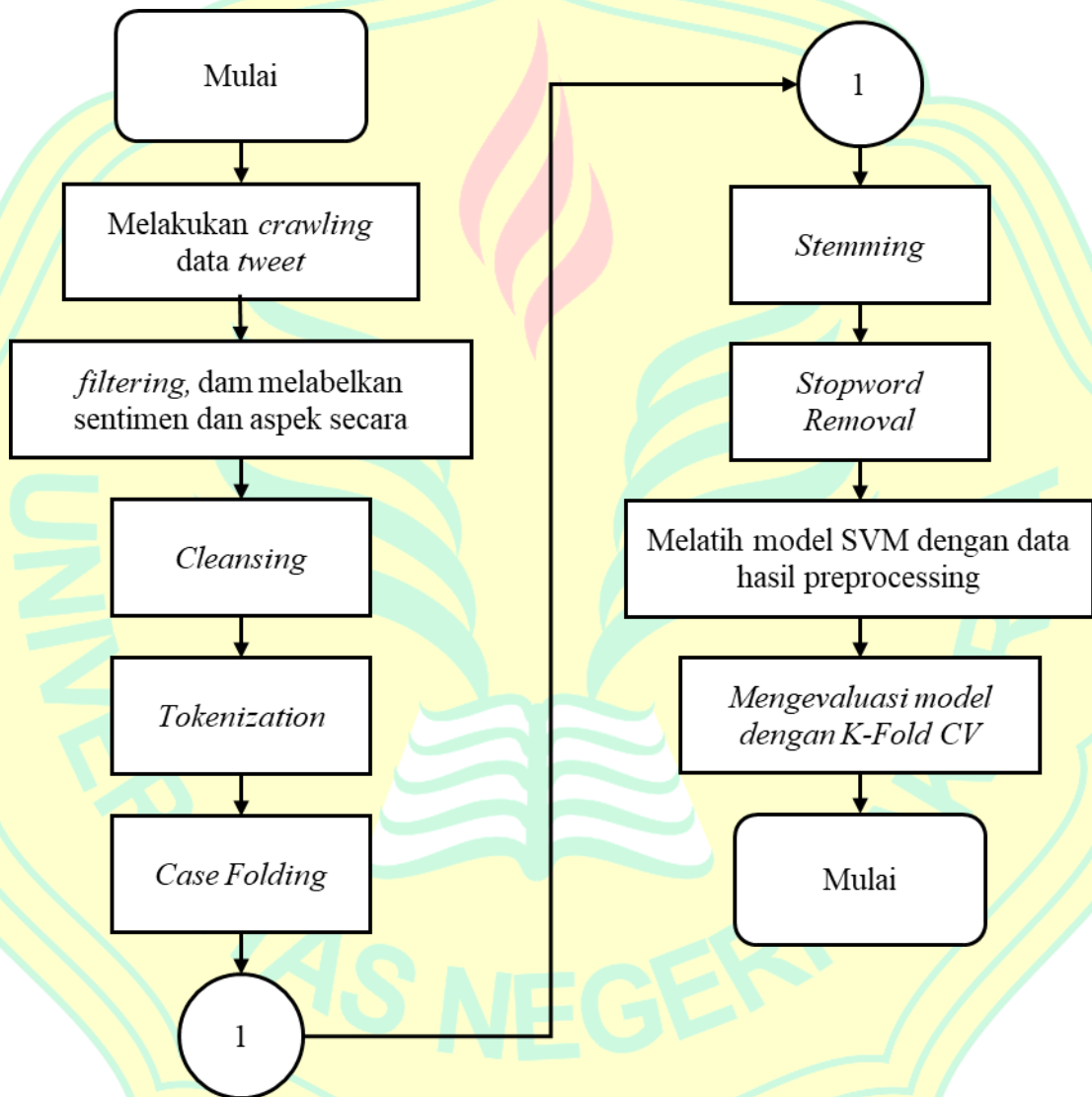
Software	Spesifikasi
Sistem Operasi (OS)	Windows 10 Home 64-bit
Driver	AMD Radeon Software Crimson ReLive Edition
Bahasa Pemrograman	Python 3.9.4
IDE	Google Colab

3.2.2 Bahan

Penelitian ini akan menggunakan data yang bersumber dari tweet dari pengguna Twitter atau yang sekarang namanya menjadi “X”. data yang di ambil hanya yang menggunakan bahasa indonesia dan diambil sejak 25 April 2023 hingga 18 Mei 2023, rentang tanggal tersebut diambil sesuai dengan keberangkatan dan kepulangan atlet Indonesia dari acara SEA Games 2023 di Kamboja. Kemudian kata kunci yang akan digunakan sebagai *Keyword* pengambilan data adalah “SEA Games 2023”, “Fasilitas SEA Games”, “SEA Games Kamboja”, “#SEAGames2023”, “#FasilitasSEAGames”

3.3 Diagram Alir Penelitian

Diagram Alir Penelitian dari proses penelitian ini dapat dilihat pada Gambar 3.1 yang memperlihatkan bagaimana proses penelitian ini akan berlangsung dari tahap awal proses pengambilan data hingga selesai.



Gambar 3.1 Flowchart Penelitian

3.3.1 Melakukan *Crawling* untuk pengambilan data *tweet*

Tahap awal dalam proses mencari sentimen berbasis aspek adalah dengan pengumpulan data menggunakan teknik *crawling*. Penelitian ini akan melakukan *crawling* dibantu dengan *library Tweepy* untuk mempermudah pengambilan data. Penggunaan *Tweepy* sebagai alat *Crawling* dapat mempermudah proses pengumpulan data di *platform* Twitter karena dapat menggunakan *Twitter API*. Kemudian data yang di ambil akan menggunakan keyword seperti “SEA Games 2023”, “Fasilitas SEA Games”, “SEA Games Kamboja”, “#SEAGames2023”, dan “#FasilitasSEAGames” sebagai *keyword* utama.

3.3.2 Mengidentifikasi Aspek pada data secara manual

Setelah data yang diambil sudah selesai, tahap selanjutnya adalah mengidentifikasi aspek pada data yang sudah tersedia. Aspek yang perlu diidentifikasi adalah aspek kebersihan, aspek kelengkapan, dan aspek kenyamanan. Data yang tidak memiliki hubungan yang relevan dengan ketiga aspek tersebut akan dihapuskan. Proses ini akan dilakukan secara manual untuk mendapatkan hasil data yang lebih akurat.

3.3.3 Melabelkan data secara manual

Selanjutnya akan dilakukan proses pelabelan data yang dilakukan secara manual untuk nantinya dimodelkan dengan menggunakan *Support Vector Machine*. Data yang diberikan label terdiri dari “1” untuk data sentimen positif, dan “-1” untuk data sentimen negatif. Contoh dapat dilihat pada Tabel 3.4.

Tabel 3.4 Contoh pelabelan manual

Opini	Aspek	Sentimen
“Ruang tunggu-nya banyak sampah berserakan”	Kebersihan	-1
“Kamar tidurnya nyaman sekali”	Kenyamanan	1
“Fasilitas gym di Wisma Atlet SEA Games 2023 Kamboja cukup lengkap dan modern. #SEAGames2023”	Kelengkapan	1

3.3.4 Preprocessing

Ditahap ini data akan diproses untuk disederhanakan, dan juga untuk mempermudah untuk nantinya dalam proses pemodelan dengan SVM. Berikut merupakan contoh penerapan *preprocessing* menggunakan contoh *tweet* berikut ini:

“Harap maklum, wong kursi aja pake korsi arisan, lampu aja pake lampu sorot mobil, sound aja pke sound senam SKJ 🤔”

1. *Cleansing*

Di tahap ini setiap simbol-simbol maupun emoji pada data akan di hapuskan karena pada penelitian ini hanya berfokus pada sentimen berbentuk teks saja sehingga simbol dan emoji dirasa tidak diperlukan. Contoh pada Tabel 3.4.

Tabel 3.5 *Cleansing*

Tweet	Cleansing
harap maklum, wong kursi aja pake korsi arisan, lampu aja pake lampu sorot mobil, sound aja pke sound senam SKJ 🤔	harap maklum wong kursi aja pake korsi arisan lampu aja pake lampu sorot mobil sound aja pke sound senam SKJ

Dari contoh Tabel 3.4 dapat dilihat setelah proses *cleansing*, tanda baca seperti koma “;” dan emoji “🤔” dihapuskan dari data yang akan di gunakan.

2. Tokenization

Pada tahap ini setiap kalimat akan dipecah menjadi kata-kata atau yang bisa disebut juga sebagai “token”. Berikut merupakan contoh penerapan *tokenization* pada tabel 3.5.

Tabel 3.6 Tokenization

Cleansing	Tokenization
harap maklum wong kursi aja pake korsi arisan lampu aja pake lampu sorot mobil sound aja pke sound senam SKJ	['harap', 'maklum', 'wong', 'kursi', 'aja', 'pake', 'korsi', 'arisan', 'lampu', 'aja', 'pake', 'lampu', 'sorot', 'mobil', 'sound', 'aja', 'pke', 'sound', 'senam', 'SKJ']

Terlihat bahwa data yang sebelumnya berbentuk teks pada umumnya berubah menjadi token-token

3. Case Folding

Kemudian terdapat tahap *case folding*. Di tahap ini setiap kata-kata yang mengandung huruf kapital (*uppercase*) akan diubah menjadi huruf kecil (*lowercase*) seperti yang terlihat pada Tabel 3.6. Terlihat bahwa token ‘SKJ’ berubah menjadi ‘skj’.

Tabel 3.7 Case Folding

Tokenization	Case Folding
['harap', 'maklum', 'wong', 'kursi', 'aja', 'pake', 'korsi', 'arisan', 'lampu', 'aja', 'pake', 'lampu', 'sorot', 'mobil', 'sound', 'aja', 'pke', 'sound', 'senam', 'SKJ']	['harap', 'maklum', 'wong', 'kursi', 'aja', 'pake', 'korsi', 'arisan', 'lampu', 'aja', 'pake', 'lampu', 'sorot', 'mobil', 'sound', 'aja', 'pke', 'sound', 'senam', 'skj']

4. *Stemming*

Proses berikutnya terdapat *Stemming*. Yaitu merubah kata-kata kebentuk dasarnya. Dan karena penelitian ini mengambil data berbahasa indonesia, maka tiap kata asing akan di *translate* kedalam bahasa indonesia dasar. Seperti yang dicontohkan pada Tabel 3.7 kata “wong” dirubah kedalam bahasa indonesia dasar yang berbarti “orang”, kata “aja” berubah menjadi “saja”, dan seterusnya.

Tabel 3.8 Stemming

Case Folding	Stemming
['harap', 'maklum', 'wong', 'kursi', 'aja', 'pake', 'korsi', 'arisan', 'lampu', 'aja', 'pake', 'lampu', 'sorot', 'mobil', 'sound', 'aja', 'pke', 'sound', 'senam', 'skj']	['harap', 'maklum', 'orang', 'kursi', 'saja', 'pakai', 'kursi', 'arisan', 'lampu', 'saja', 'pakai', 'lampu', 'sorot', 'mobil', 'suara', 'saja', 'pakai', 'suara', 'senam', 'skj']

5. *Stopword Removal*

Tahap terakhir adalah *Stopword Removal*, yaitu penghapusan kata-kata yang dianggap umum dan tidak memberikan banyak informasi dalam analisis teks. Dapat dilihat pada contoh Tabel 3.8, token ‘harap’, ‘maklum’, ‘orang’, ‘saja’, dan ‘pakai’ dihapuskan dalam teks yang sudah melalau proses *Stopword Removal*.

Tabel 3.9 Stopword Removal

Stemming	Stopword Removal
['harap', 'maklum', 'orang', 'kursi', 'saja', 'pakai', 'kursi', 'arisan', 'lampu', 'saja', 'pakai', 'lampu', 'sorot', 'mobil', 'suara', 'saja', 'pakai', 'suara', 'senam', 'skj']	['orang', 'kursi', 'arisan', 'lampu', 'sorot', 'mobil', 'suara', 'senam', 'skj']

3.3.5 Menganalisis sentimen dengan SVM

Tahap selanjutnya setelah dilakukan *preprocessing* dan pelabelan sentimen adalah menganalisis sentimen. Proses ini akan dilakukan pada *web app* Google Colab menggunakan bahasa Python dan algoritma *Support Vector Machine* untuk dijadikan model. Dikarenakan hasil yang ingin dicapai hanya sentimen positif dan negatif pada setiap aspek, maka proses klasifikasi akan menggunakan metode klasifikasi linear.

3.3.6 Evaluasi hasil dengan K-Folds CV

Proses terakhir adalah mengevaluasi model yang sudah berhasil di latih dengan menggunakan metode K-Folds Cross Validation. Fold yang akan digunakan dalam penelitian adalah 10 Fold. Contoh penggunaan K-Folds Cross Validation dengan 10 Dataset dan K = 10 adalah sebagai berikut:

Fold 1 = A
Fold 2 = B
Fold 3 = C
Fold 4 = D
Fold 5 = E

Fold 6 = F
Fold 7 = G
Fold 8 = H
Fold 9 = I
Fold 10 = J

Tabel 3.10 Contoh K-fold Cross Validation

Fold	Data Latih	Data Uji
1	Fold 2, Fold 3, Fold 4, Fold 5, Fold 6, Fold 7, Fold 8, Fold 9, Fold 10	Fold 1
2	Fold 1, Fold 3, Fold 4, Fold 5, Fold 6, Fold 7, Fold 8, Fold 9, Fold 10	Fold 2
3	Fold 1, Fold 2, Fold 4, Fold 5, Fold 6, Fold 7, Fold 8, Fold 9, Fold 10	Fold 3
4	Fold 1, Fold 2, Fold 3, Fold 5, Fold 6, Fold 7, Fold 8, Fold 9, Fold 10	Fold 4
5	Fold 1, Fold 2, Fold 3, Fold 4, Fold 6, Fold 7, Fold 8, Fold 9, Fold 10	Fold 5
6	Fold 1, Fold 2, Fold 3, Fold 4, Fold 5, Fold 7, Fold 8, Fold 9, Fold 10	Fold 6
7	Fold 1, Fold 2, Fold 3, Fold 4, Fold 5, Fold 6, Fold 8, Fold 9, Fold 10	Fold 7
8	Fold 1, Fold 2, Fold 3, Fold 4, Fold 5, Fold 6, Fold 7, Fold 9, Fold 10	Fold 8
9	Fold 1, Fold 2, Fold 3, Fold 4, Fold 5, Fold 6, Fold 7, Fold 8, Fold 10	Fold 9
10	Fold 1, Fold 2, Fold 3, Fold 4, Fold 5, Fold 6, Fold 7, Fold 8, Fold 9	Fold 10

Setelah mendapatkan nilai akurasi dari masing-masing fold, langkah selanjutnya adalah menghitung akurasi rata-rata dengan rumus (10):

$$akurasi\ rata - rata = \frac{\sum akurasi\ k}{k}$$

3.4 Teknik dan Prosedur Pengumpulan Data

Berdasarkan Gambar 3.1 Teknik yang akan digunakan untuk pengumpulan data adalah dengan menggunakan *Crawling*. Dengan menggunakan teknik *crawling*, data-data yang di butuh dalam *platform* Twitter dapat diambil. Namun untuk menggunakan teknik tersebut perlu sebuah alat/*library* khusus salah satunya

adalah *Tweepy*. Dengan bantuan *library* *Tweepy* data yang akan diambil dari *platform* Twitter adalah sebagai berikut:

- *Date*/Tanggal dari tweet ketika diunggah
- *Username*/ID dari tweet pengguna Twitter
- Isi tweet yang didalamnya memiliki tidak lebih dari 300 karakter

Data yang sudah di *Crawling* kemudian akan di ubah menjadi format *CSV*. Hal ini dilakukan untuk mempermudah proses pelabelan dan proses *Preprocessing* nantinya. Dimana proses pelabelan akan dilakukan secara manual dengan mengidentifikasi aspek kenyamanan, kelengkapan, dan kebersihan. Setelah dilakukan pelabelan, data tersebut akan dilakukan *Preprocessing* untuk mempermudah pemodelan menggunakan metode *Support Vector Machine*.

3.5 Teknik Analisis Data

Setelah proses pemodelan selesai, berikutnya model akan di evaluasi dengan menggunakan metode *K-Folds Cross Validation*. Dimana data sebelumnya akan dikelompokkan menjadi data *Training* dan data *Testing*. Jumlah pengelompokan bergantung pada jumlah *Folds* yang akan di gunakan, dalam penelitian ini *Folds* yang akan di gunakan adalah $K = 10$. Setelah dilakukan perhitungan akurasi pada masing-masing *Fold*, selanjutnya Akurasi akan di jumlahkan dan dibagi dengan jumlah *Fold* untuk menemukan akurasi rata-rata dari model yang sudah dibuat dengan rumus:

$$\text{akurasi rata - rata} = \frac{\sum \text{akurasi } k}{k}$$

Dari metode inilah dapat diketahui seberapa baik akurasi rata-rata dari model *machine learning* yang sudah dibuat.

BAB IV

HASIL PENELITIAN

4.1 Pengumpulan Data

Penelitian ini dilakukan untuk mendapatkan gambaran umum tentang sentimen warga Twitter serta mengetahui seberapa besar nilai akurasi model machine learning dengan algoritma *Support Vector Machine* pada opini setiap warga Twitter tentang Fasilitas SEA Games 2023.

Penelitian ini dilakukan dengan menarik sejumlah besar data opini warga Twitter dengan bantuan Tweepy yang dijalankan pada IDE milik google yaitu Google Colab. Penelitian ini menggunakan *keyword* “SEA Games”, “Fasilitas”, “Kamar atlit”, “Lampu Mobil” dan/atau “Kamar Ganti” sebagai kata kunci utama dalam pencarian sumber data penelitian.

username	created_at	full_text
nagabolasports	Wed May 31 23:29:15	Erick Thohir lalu memaparkan kalender Timnas Indo
mataberita_	Wed May 31 23:26:26	Rektor USM Apresiasi Prestasi Dewangga Mahasiswa
kompasiana	Wed May 31 17:21:34	Kontroversi SEA Games 2023: Ketidaksetaraan Perlak
news24_coid	Wed May 31 17:04:14	Golf House Srixon Beri Penghargaan Ke Peraih Emas
ResNgawi	Wed May 31 17:01:30	Apresiasi Kapolres Ngawi AKBP Dwiasi Wiyatputera
Myeshaa27	Wed May 31 15:36:17	Prestasi @erickthohir semakin mencuat ketika Timn
radarbangsa_com	Wed May 31 14:58:03	Komisi X DPR RI Apresiasi Kontingen Indonesia Rebu
KompasTV	Wed May 31 14:54:34	AVC Challenge Cup 2023: Dihuni Mayoritas Skwad SEA
Kapolreslmy	Wed May 31 14:53:10	Hari ini saya bersilahturahmi dengan Mba Leni seora

Gambar 4.1 Hasil Crawling Data Twitter dengan Tweepy

Gambar 4.1 merupakan hasil penarikan data (*crawling*). Dari hasil *crawling* didapatkan data sebanyak 1994 data mentah yang belum di *filter*. Setelah mendapatkan data yang sudah di *crawling*, data tersebut kemudian dibuat menjadi berformat file .CSV yang kemudian dipindahkan ke dalam excel untuk dilakukan proses *filtering*, dan pelabelan manual.

Username	Waktu	Teks
MustBeeajah	Mon May 29 02:32:	Kemenangan Timnas sepak bola U-22 Indonesia di SEA GAMES
Herikurrrrr	Mon May 29 02:10:	Semenjak di era kepemimpinan Ketua PSSI @erickthohir waja
GemoyBahagia	Mon May 29 02:01:	Kamar tidurnya luas dan nyaman. #SEAGames2023
Syarifa_hh	Mon May 29 01:54:	Fasilitas di Wisma Atlet SEA Games 2023 Kamboja cukup lengk
hanah_ayu	Mon May 29 01:34:	Fasilitas di Wisma Atlet SEA Games 2023 Kamboja cukup bersi
BUMN_BDG_Beda	Mon May 29 01:31:	Lapangan latihan untuk sepak bola di SEA Games 2023 Kambo
AsriKamilaa	Mon May 29 01:30:	Lapangan latihan untuk sepak bola di SEA Games 2023 Kambo
golkarpedia	Mon May 29 01:30:	Kolam renang di SEA Games 2023 Kamboja terlihat bersih dan
Aqilanews	Mon May 29 01:16:	Kantin di Wisma Atlet SEA Games 2023 Kamboja menyediakan

Gambar 4.2 Hasil dari *filtering*

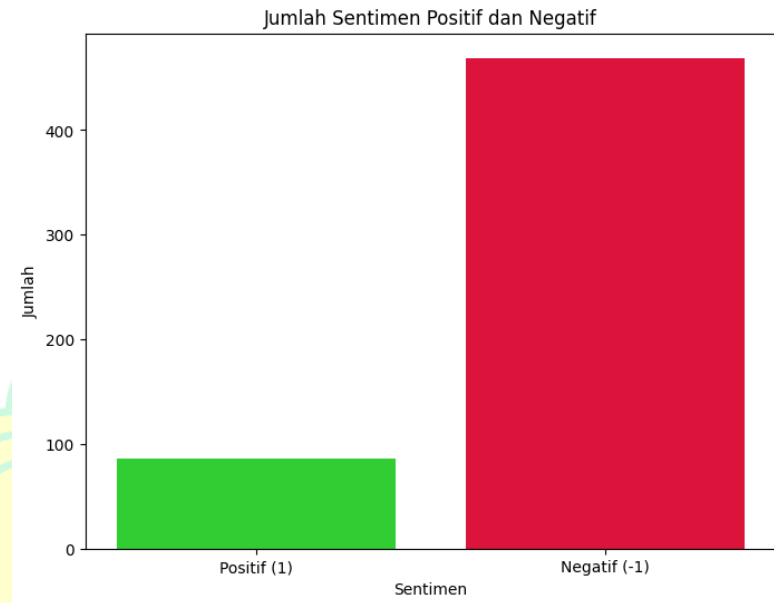
Pada Gambar 4.2 diatas, merupakan data yang berhasil disaring. Data tersebut berkurang menjadi 555 data untuk menyesuaikan kebutuhan kriteria penelitian. kriteria tersebut adalah opini-opini warga Twitter yang memiliki keterkaitannya dengan aspek “kebersihan”, “kelengkapan”, dan “kenyamanan”.

Data kemudian melalui proses pelabelan secara manual. Data yang memiliki sentimen positif akan diberi simbol “1” dan data yang memiliki sentimen negatif akan diberi simbol “-1”. Data juga melalui proses klasifikasi aspek yang terdiri dari “kebersihan”, “kelengkapan”, dan “kenyamanan”. Gambar 4.3 akan memperlihatkan bagaimana hasil dari proses pelabelan manual..

sikeu_batola	Sat May 27 23:08:3	Saya baru saja mencoba makanan di kafetaria SEA Games. Ma	1 Kelengkap
Dutajomblo3	Sat May 27 23:07:1	Saya merasa sangat nyaman selama berada di wisma atlet SE	1 Kenyamar
PolsekBarambai	Sat May 27 22:57:1	@douciz Nyaman bgt	1 Kenyamar
ReskrimBatola	Sat May 27 22:30:1	@Reza84705650 Kak Reza maaf ya udah bikin ga nyaman. Peril	-1 Kenyamar
zebrabatola	Sat May 27 22:29:4	@dimasmisme @Seressahh @idextratime @Santi_J_FM Ngapa	1 Kelengkap
PolsekBarambai	Sat May 27 22:19:5	Kemeja Bersih Wanita Drape Longgar bahan adem halus dikul	1 Kebersiha
JatimPemprov	Sat May 27 19:07:0	@adith_wp Aku sampe belain ke iBox Central Park buat purcha	1 Kelengkap
cakranewsonline	Sat May 27 19:04:3	@lucaneshiro Cucinya kurang bersih yaa	-1 Kebersiha
sektambancity	Sat May 27 19:04:3	@lampungbase Cerita itu yg lengkap nder kek gini kronologi ny	-1 Kelengkap
Satnarkobabt	Sat May 27 17:07:5	@kochengfs Kucingku kalo ketakutan berak. Itu udah ngeong t	-1 Kenyamar

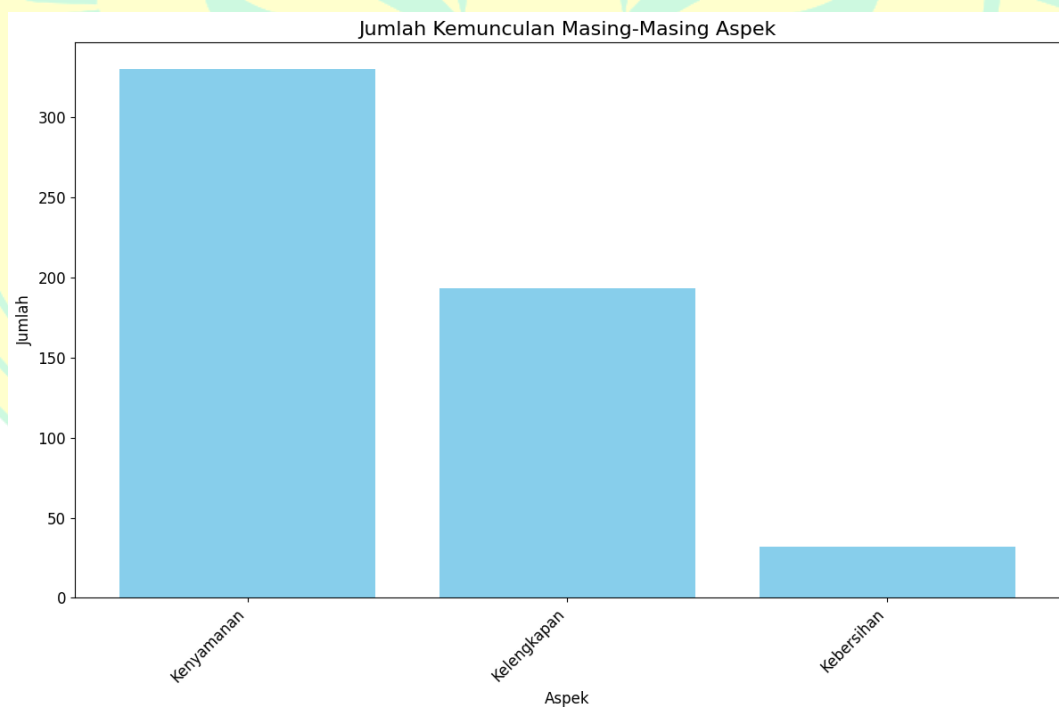
Gambar 4.3 Data setelah pelabelan manual

Berdasarkan Gambar 4.3 terdapat 6 opini warga twitter yang sentimen-nya bernilai positif yang disimbolkan dengan “1” , kemudian terdapat 4 opini yang bernilai sentimen negatif yang diberi simbol “-1”. Pada total 555 data yang sudah labeling, terdapat 86 opini yang diberi sentimen positif dan 469 opini yang diberi sentimen negatif.



Gambar 4.4 Gambar Perbandingan data positif dan negatif

Dari data yang sudah dilabelkan secara manual terdapat juga label aspek berupa “kebersihan”, “kelengkapan” dan “kenyamanan”. Gambar 4.5 akan memperlihatkan perbandingan jumlah setiap aspek yang ada diantara 555 data.



Gambar 4.5 Perbandingan jumlah aspek

Dari gambar 4.5 dapat terlihat kemunculan opini dengan aspek kenyamanan, muncul sebagai opini yang paling banyak yaitu sebesar 330 data, aspek kelengkapan berjumlah 193 data, dan aspek kebersihan berjumlah 32 data.

4.2 Preprocessing Data

Setelah data dilabelkan secara manual. Data kemudian memasuki tahapan *preprocessing*, tahapan ini dilakukan agar proses pemodelan machine learning dapat berjalan lebih efisien dan akurat.

4.2.1 Cleansing

Cleansing merupakan proses pembersihan data dengan menghilangkan *noise* yang dapat mengganggu keakuratan data. *Cleansing* bekerja dengan menghilangkan *emation*, angka, tanda baca, *link*, dan *whitespace*. Proses *Cleansing* dilakukan dengan bantuan modul *regular expression*. Tabel 4.1 merupakan contoh dari penggunaan *Cleansing* pada data yang sudah dilabelkan.

Tabel 4.1 Perbandingan setelah di Cleansing

<i>Tweet</i>	<i>Cleansing</i>
Tak Ada Lampu sampai Kursi Hajatan SEA Games 2023 Kamboja Tak Siap? Penyelenggaraan SEA Games 2023 di Kamboja mendapat sorotan netizen Tanah Air dikarenakan banyak kejadian konyol yang dialami karena ketidaksiapan panitia. https://t.co/1MCsokwkRr	Tak Ada Lampu sampai Kursi Hajatan SEA Games Kamboja Tak Siap Penyelenggaraan SEA Games di Kamboja mendapat sorotan netizen Tanah Air dikarenakan banyak kejadian konyol yang dialami karena ketidaksiapan panitia
@Thitipan_Mafia @mazzini_gsp ini sea games jelek banget lebih jelek daripada sea games Filipina kmrin. Mulai dari podium pake lampu mobil atlet impor wasit gaje ngancam atlet fasilitas seadanya dll. Buat yang bela Kamboja karena mereka negara miskin gak negara miskin gak gini. Ini namanya ngawur	ini sea games jelek banget lebih jelek daripada sea games Filipina kmrin. Mulai dari podium pake lampu mobil atlet impor wasit gaje ngancam atlet fasilitas seadanya dll. Buat yang bela Kamboja karena mereka negara miskin gak negara miskin gak gini. Ini namanya ngawur

4.2.2 Tokenization

Setelah dilakukan proses *cleansing*, data kemudian melalui proses *Tokenization*. Proses ini dilakukan dengan mengubah teks *tweet* menjadi Token per kata. Dengan bantuan *module NLTK* didapatkan hasil *Tokenization* seperti yang diperlihatkan pada Tabel 4.2.

Tabel 4.2 Perbandingan data Cleansing dengan Tokenization

<i>Tweet Cleansing</i>	<i>Tokenization</i>
Tak Ada Lampu sampai Kursi Hajatan SEA Games Kamboja Tak Siap Penyelenggaraan SEA Games di Kamboja mendapat sorotan netizen Tanah Air dikarenakan banyak kejadian konyol yang dialami karena ketidaksiapan panitia	['Tak', 'Ada', 'Lampu', 'sampai', 'Kursi', 'Hajatan', 'SEA', 'Games', 'Kamboja', 'Tak', 'Siap', 'Penyelenggaraan', 'SEA', 'Games', 'di', 'Kamboja', 'mendapat', 'sorotan', 'netizen', 'Tanah', 'Air', 'dikarenakan', 'banyak', 'kejadian', 'konyol', 'yang', 'dialami', 'karena', 'ketidaksiapan', 'panitia']
ini sea games jelek banget lebih jelek daripada sea games Filipina kmrin Mulai dari podium pake lampu mobil atlet impor wasit gaje ngancam atlet fasilitas seadanya dll Buat yang bela Kamboja karena mereka negara miskin gak negara miskin gak gini Ini namanya ngawur	['ini', 'sea', 'games', 'jelek', 'banget', 'lebih', 'jelek', 'daripada', 'sea', 'games', 'Filipina', 'kmrin', 'Mulai', 'dari', 'podium', 'pake', 'lampu', 'mobil', 'atlet', 'impor', 'wasit', 'gaje', 'ngancam', 'atlet', 'fasilitas', 'seadanya', 'dll', 'Buat', 'yang', 'bela', 'Kamboja', 'karena', 'mereka', 'negara', 'miskin', 'gak', 'negara', 'miskin', 'gak', 'gini', 'Ini', 'namanya', 'ngawur']

4.2.3 Case Folding

Setelah data sudah menjadi *Token*, data kemudian dilakukan proses *Case Folding* atau mengubah semua *Token* kata menjadi huruf kecil (*lowercase*), proses ini dilakukan untuk mempermudah pengolahan. Tabel 4.3 akan memperlihatkan perbandingannya.

Tabel 4.3 Perbandingan data Tokenization dan Case Folding

<i>Tweet Tokenization</i>	<i>Case Folding</i>
['Tak', 'Ada', 'Lampu', 'sampai', 'Kursi', 'Hajatan', 'SEA', 'Games', 'Kamboja', 'Tak', 'Siap', 'Penyelenggaraan', 'SEA', 'Games', 'di', 'Kamboja', 'mendapat', 'sorotan', 'netizen', 'Tanah', 'Air', 'dikarenakan', 'banyak', 'kejadian', 'konyol', 'yang', 'dialami', 'karena', 'ketidaksiapan', 'panitia']	['tak', 'ada', 'lampu', 'sampai', 'kursi', 'hajatan', 'sea', 'games', 'kamboja', 'tak', 'siap', 'penyelenggaraan', 'sea', 'games', 'di', 'kamboja', 'mendapat', 'sorotan', 'netizen', 'tanah', 'air', 'dikarenakan', 'banyak', 'kejadian', 'konyol', 'yang', 'dialami', 'karena', 'ketidaksiapan', 'panitia']
['ini', 'sea', 'games', 'jelek', 'banget', 'lebih', 'jelek', 'daripada', 'sea', 'games', 'Filipina', 'kmrin', 'Mulai', 'dari', 'podium', 'pake', 'lampu', 'mobil', 'atlet', 'impor', 'wasit', 'gaje', 'ngancam', 'atlet', 'fasilitas', 'seadanya', 'dll', 'Buat', 'yang',	['ini', 'sea', 'games', 'jelek', 'banget', 'lebih', 'jelek', 'daripada', 'sea', 'games', 'filipina', 'kmrin', 'mulai', 'dari', 'podium', 'pake', 'lampu', 'mobil', 'atlet', 'impor', 'wasit', 'gaje', 'ngancam', 'atlet', 'fasilitas', 'seadanya', 'dll', 'buat', 'yang',

'bela', 'Kamboja', 'karena', 'mereka', 'negara', 'miskin', 'gak', 'negara', 'miskin', 'gak', 'gini', 'Ini', 'namanya', 'ngawur']	'bela', 'kamboja', 'karena', 'mereka', 'negara', 'miskin', 'gak', 'negara', 'miskin', 'gak', 'gini', 'ini', 'namanya', 'ngawur']
--	--

4.2.4 Stemming

Stemming adalah proses untuk mengurangi semua kata dengan batang kata yang sama menjadi bentuk umum. Tujuan utama *stemming* adalah untuk meningkatkan efisiensi pencarian informasi dengan mengurangi jumlah kata yang harus diindeks dan dicari. Dengan bantuan module sastrawi di dapatkan hasil *stemming* seperti yang ditunjukkan pada Tabel 4.4 dibawah ini.

Tabel 4.4 Perbandingan Data Case Folding dan Stemming

<i>Tweet Case Folding</i>	<i>Stemming</i>
['tak', 'ada', 'lampu', 'sampai', 'kursi', 'hajatan', 'sea', 'games', 'kamboja', 'tak', 'siap', 'penyelenggaraan', 'sea', 'games', 'di', 'kamboja', 'mendapat', 'sorotan', 'netizen', 'tanah', 'air', 'dikarenakan', 'banyak', 'kejadian', 'konyol', 'yang', 'dialami', 'karena', 'ketidaksiapan', 'panitia']	['tak', 'ada', 'lampu', 'sampai', 'kursi', 'hajatan', 'sea', 'games', 'kamboja', 'tak', 'siap', 'selenggara', 'sea', 'games', 'di', 'kamboja', 'dapat', 'sorot', 'netizen', 'tanah', 'air', 'karena', 'banyak', 'jadi', 'konyol', 'yang', 'alami', 'karena', 'ketidaksiapan', 'panitia']
['ini', 'sea', 'games', 'jelek', 'banget', 'lebih', 'jelek', 'daripada', 'sea', 'games', 'filipina', 'kmrin', 'mulai', 'dari', 'podium', 'pake', 'lampu', 'mobil', 'atlet', 'impor', 'wasit', 'gaje', 'ngancam', 'atlet', 'fasilitas', 'seadanya', 'dll', 'buat', 'yang', 'bela', 'kamboja', 'karena', 'mereka', 'negara', 'miskin', 'gak', 'negara', 'miskin', 'gak', 'gini', 'ini', 'namanya', 'ngawur']	['ini', 'sea', 'games', 'jelek', 'banget', 'lebih', 'jelek', 'daripada', 'sea', 'games', 'filipina', 'kmrin', 'mulai', 'dari', 'podium', 'pake', 'lampu', 'mobil', 'atlet', 'impor', 'wasit', 'gaje', 'ngancam', 'atlet', 'fasilitas', 'ada', 'dll', 'buat', 'yang', 'bela', 'kamboja', 'karena', 'mereka', 'negara', 'miskin', 'gak', 'negara', 'miskin', 'gak', 'gini', 'ini', 'nama', 'ngawur']

4.2.5 Stopword Removal

Stopword Removal yaitu penghapusan kata-kata yang dianggap umum dan tidak memberikan banyak informasi. Tujuannya untuk mengurangi jumlah data yang disimpan dalam *Token*. Proses *Stopword Removal* berlangsung dengan bantuan modul Sastrawi. Tabel 4.5 akan memperlihatkan perbandingannya.

Tabel 4.5 Perbandingan data Stemming dan Stopword Removal

<i>Tweet Stemming</i>	<i>Stopword Removal</i>
['tak', 'ada', 'lampu', 'sampai', 'kursi', 'hajat', 'sea', 'games', 'kamboja', 'tak', 'siap', 'selenggara', 'sea', 'games', 'di', 'kamboja', 'dapat', 'sorot', 'netizen', 'tanah', 'air', 'karena', 'banyak', 'jadi', 'konyol', 'yang', 'alami', 'karena', 'ketidaksiapan', 'panitia']	['tak', 'lampu', 'kursi', 'hajat', 'sea', 'games', 'kamboja', 'tak', 'siap', 'selenggara', 'sea', 'games', 'kamboja', 'sorot', 'netizen', 'tanah', 'air', 'banyak', 'jadi', 'konyol', 'alami', 'ketidaksiapan', 'panitia']
['ini', 'sea', 'games', 'jelek', 'banget', 'lebih', 'jelek', 'daripada', 'sea', 'games', 'filipina', 'kmrin', 'mulai', 'dari', 'podium', 'pake', 'lampu', 'mobil', 'atlet', 'impor', 'wasit', 'gaje', 'ngancam', 'atlet', 'fasilitas', 'ada', 'dll', 'buat', 'yang', 'bela', 'kamboja', 'karena', 'mereka', 'negara', 'miskin', 'gak', 'negara', 'miskin', 'gak', 'gini', 'ini', 'nama', 'ngawur']	['sea', 'games', 'jelek', 'banget', 'lebih', 'jelek', 'sea', 'games', 'filipina', 'kmrin', 'mulai', 'podium', 'pake', 'lampu', 'mobil', 'atlet', 'impor', 'wasit', 'gaje', 'ngancam', 'atlet', 'fasilitas', 'buat', 'bela', 'kamboja', 'negara', 'miskin', 'gak', 'negara', 'miskin', 'gak', 'gini', 'nama', 'ngawur']

4.3 Hasil Penelitian dengan *Support Vector Machine*

Algoritma *Support Vector Machine* merupakan algoritma machine learning yang memiliki prinsip dasar *linear classifier*, yang berarti SVM dapat bekerja pada kasus klasifikasi yang secara linier dapat dipisahkan. Namun sekarang SVM dapat bekerja pada dimensi yang lebih tinggi dengan bantuan karnel. Pada penelitian ini akan menggunakan kernel linear. Kemudian data yang sudah di preprocessing akan di ubah menjadi representasi numerik berbasis TF-IDF.

4.3.1 Evaluasi Model Setiap Fold

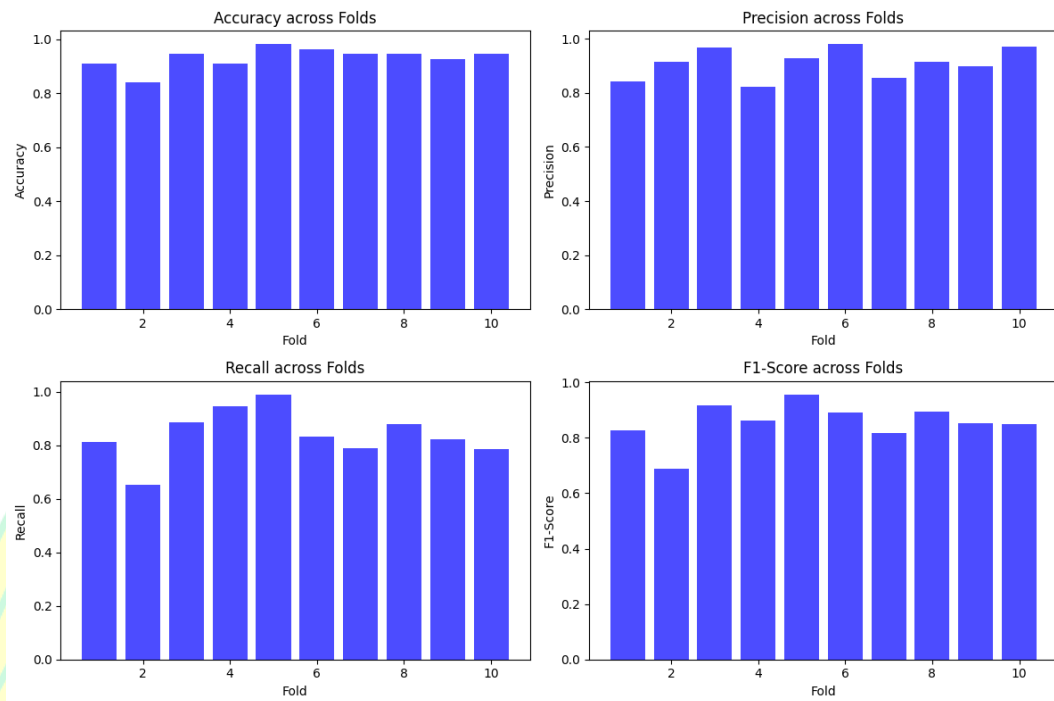
Setelah model SVM dibuat, selanjutnya model di evaluasi dengan metode K-Fold Cross Validation. Penelitian ini menggunakan K=10 dan data akan di acak terlebih dahulu sebelum dilakukan pembagian *Fold*. Kemudian di dapatkan hasil penilaian performa model machine learning dengan matriks akurasi, presisi, *recall*, dan *F1-score* setiap *Fold* sebesar:

Tabel 4. 6 Hasil evaluasi model setiap Fold

=== Fold 1 === Accuracy: 0.9107142857142857 Precision: 0.84375 Recall: 0.8120567375886525 F1-Score: 0.826625386996904	=== Fold 2 === Accuracy: 0.8392857142857143 Precision: 0.9134615384615384 Recall: 0.6538461538461539 F1-Score: 0.6879256965944271
---	---

<p>==== Fold 3 ====</p> <p>Accuracy: 0.9464285714285714 Precision: 0.9673913043478262 Recall: 0.8846153846153846 F1-Score: 0.9179286761113825</p>	<p>==== Fold 4 ====</p> <p>Accuracy: 0.9107142857142857 Precision: 0.8214285714285714 Recall: 0.9468085106382979 F1-Score: 0.8632144601856375</p>
<p>==== Fold 5 ====</p> <p>Accuracy: 0.9821428571428571 Precision: 0.9285714285714286 Recall: 0.99 F1-Score: 0.9564879564879565</p>	<p>==== Fold 6 ====</p> <p>Accuracy: 0.9636363636363636 Precision: 0.9803921568627452 Recall: 0.8333333333333333 F1-Score: 0.89</p>
<p>==== Fold 7 ====</p> <p>Accuracy: 0.9454545454545454 Precision: 0.8553921568627452 Recall: 0.79 F1-Score: 0.8184818481848184</p>	<p>==== Fold 8 ====</p> <p>Accuracy: 0.9454545454545454 Precision: 0.9162234042553192 Recall: 0.8780193236714976 F1-Score: 0.8956356736242885</p>
<p>==== Fold 9 ====</p> <p>Accuracy: 0.9272727272727272 Precision: 0.8973214285714286 Recall: 0.822463768115942 F1-Score: 0.8537234042553192</p>	<p>==== Fold 10 ====</p> <p>Accuracy: 0.9454545454545454 Precision: 0.9705882352941176 Recall: 0.7857142857142857 F1-Score: 0.8484848484848485</p>

Berdasarkan Tabel 4.6, terlihat bahwa persentase *Accuracy* terbesar terjadi ketika Fold ke-5 sebesar 98.21%, persentase *Precision* terbesar pada Fold ke-6 sebesar 98.03%, *Recall* terbesar pada *Fold* ke-5 sebesar 99%, dan *F1-Score* terbesar pada *Fold* ke-5 sebesar 95.64%. Setelah mengetahui besaran nilai matrix masing-masing *Fold*, kemudian dihitunglah rata-rata performa untuk model SVM yang sudah dibuat yaitu *Accuracy* 93.16%, *Precision* 90.94%, *Recall* 83.96%, dan *F1-Score* 85.58%.



Gambar 4.6 Grafik Hasil evaluasi modeFold

Gambar 4.6 menampilkan grafik dari masing-masing matriks akurasi setiap *Fold*. Terlihat dari Gambar 4.6, *Fold* ke-5 memiliki *accuracy*, *Recall*, dan *F1-Score* terbaik diantara *Fold* yang lainnya. Kemudian pada *Fold* ke-6 memiliki *Precision* terbaik diantara *Fold* yang lainnya.

4.3.2 Evaluasi Model Berdasarkan Aspek

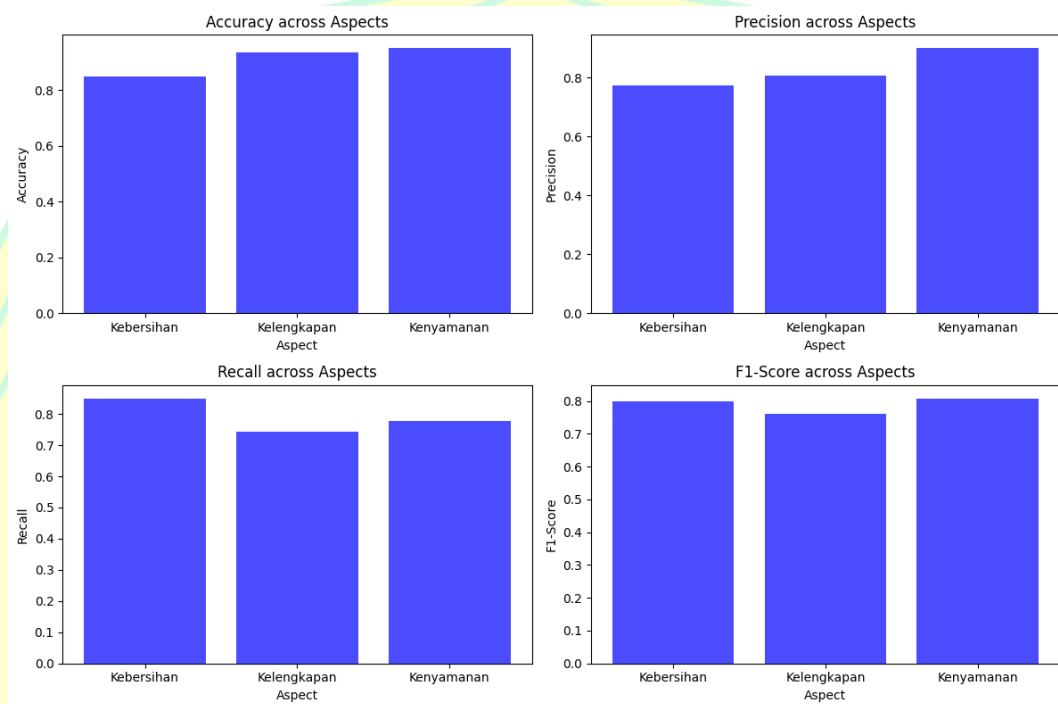
Setelah mengetahui besaran matriks pada setiap *Fold*, penelitian kemudian dilanjutkan untuk mencari tahu besaran nilai matriks untuk setiap aspek-aspek yang sudah di tentukan untuk menganalisis Sentimen Berbasis Aspek. Perhatikan Tabel 4.7 dibawah ini.

Tabel 4. 7 Matriks Hasil Evaluasi per Aspek

	Kebersihan	Kelengkapan	Kenyamanan
<i>Accuracy</i>	85%	93.73%	95.15%
<i>Precision</i>	77.5%	80.65%	90%
<i>Recall</i>	85%	74.24%	77.8%
<i>F1-Score</i>	79.83%	76.21%	80.81%

Pada Tabel 4.7 terlihat hasil evaluasi model machine learning berdasarkan aspek-aspek yang sudah di tentukan. Berdasarkan Tabel tersebut *Accuracy* tertinggi

ketika evaluasi didapatkan dari aspek Kenyamanan yang bernilai 95.15%, Nilai evaluasi *Precision* terbesar terdapat pada aspek Kenyamanan dengan nilai sebesar 90%, Nilai evaluasi *Recall* terbesar terdapat pada aspek Kebersihan dengan nilai 85%, Nilai *F1-Score* terbesar terdapat pada aspek Kenyamanan dengan nilai sebesar 80.81%.



Gambar 4.7 Grafik evaluasi per aspek

4.4 Pembahasan

Dengan menggunakan 555 data yang sudah di *filter* menurut kriteria penelitian yaitu aspek “kebersihan”, “kelengkapan”, dan “kenyamanan”. Di dapatkan hasil 86 *Tweet* yang diberi sentimen positif dan 469 *Tweet* yang diberi sentimen negatif. Kemudian didapatkan juga *Tweet* yang mengandung aspek kebersihan sebanyak 32 *Tweet*, aspek kelengkapan berjumlah 193 *Tweet*, dan aspek kenyamanan berjumlah 330 *Tweet*.

Dataset kemudian memasuki proses *preprocessing* yang bertujuan menghilangkan *noise* dan mempermudah pembuatan model machine learning. Proses yang dilewati meliputi proses *Cleansing*, *Tokenization*, *Case Folding*, *Stemming*, dan *Stopword Removal*. Dengan melalui semua langkah ini, dataset yang awalnya tidak terstruktur dan berisik diubah menjadi data yang lebih bersih,

terstruktur, dan siap untuk digunakan dalam pembuatan model machine learning yang dapat diandalkan.

Selanjutnya data yang sudah melalui proses *Preprocessing* akan ditransformasi menjadi representasi numerik berbasis TF-IDF. Setelah proses transformasi, dataset kemudian digunakan untuk proses pelatihan model *machine learning*. *Support Vector Machine* dengan kernel linear.

Setelah pemodelan berhasil dibangun, berikutnya model di evaluasi dengan metode K-Fold Cross Validation dengan menghitung nilai matriks dari *Accuracy*, *precision*, *Recall*, dan *F1-Score*. Penelitian pertama dilakukan dengan tujuan untuk mengetahui seberapa baik nilai matriks model terhadap setiap Fold jika dilakukan sebanyak K=10. Di dapatkan hasil *Accuracy*, *Recall*, *F1-Score* terbaik terjadi ketika Fold ke-5 sebesar 98.21%, 99%, dan 95.64%. Kemudian untuk persentase Precision terbaik terjadi pada Fold ke-6 sebesar 98.03%.

Penelitian selanjutnya dilakukan untuk mengetahui seberapa baik model terhadap aspek yang sudah ditentukan dengan metode K-Fold CV, dan didapatkan hasil *Accuracy* terbaik didapatkan dari aspek Kenyamanan yang bernilai 95.15%, Nilai *Precision* terbaik terdapat pada aspek Kenyamanan dengan nilai 90%, Nilai *Recall* terbaik terdapat pada aspek Kebersihan dengan nilai 85%%, Nilai *F1-Score* terbaik terdapat pada aspek Kenyamanan dengan nilai sebesar 80.81%.

4.5 Aplikasi Hasil Penelitian

Penelitian ini bertujuan untuk mendapatkan nilai akurasi algoritma *Support Vector Machine* pada kasus analisis sentimen berbasis aspek pada *Tweet* masyarakat di Twitter tentang Fasilitas SEA Games dan dapat mendeskripsikan *Tweet* pengguna Twitter yang lebih banyak mengemukakan pendapat negatif dari pada pendapat positif tentang Fasilitas yang diberikan oleh panitia SEA Games 2023. Hasil yang didapatkan dapat diimplementasikan untuk analisis sentimen berbasis aspek selanjutnya pada topik yang serupa dan dengan algoritma yang sama.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan penelitian pada BAB IV dapat disimpulkan beberapa point-point penting seperti berikut:

- 1 Setelah dilakukakn penarikan data pada *tweet* dari warga Twitter yang membahas tentang fasilitas SEA Games 2023. Penelitian ini mendapatkan 1994 data mentah. Data tersebut diolah kembali dan menjadi 555 data. Dataset kemudian dilabelkan secara manual dan didapatkan hasil 86 *Tweet* yang diberi sentimen positif dan 469 *Tweet* yang diberi sentimen negatif. Kemudian didapatkan juga *Tweet* yang mengandung aspek kebersihan sebanyak 32 *Tweet*, aspek kelengkapan berjumlah 193 *Tweet*, dan aspek kenyamanan berjumlah 330 *Tweet*.
- 2 Setelah pemodelan machine learning dengan algoritma SVM, dilakukan K-fold CV. Penelitian pertama dilakukan untuk mengetahui seberapa baik model melihat dari setiap fold yang dijalankan. Kemudian di dapatkan hasil *Accuracy*, *Recall*, *F1-Score* terbaik terjadi ketika Fold ke-5 sebesar 98.21%, 99%, dan 95.64%. Kemudian untuk persentase *Precision* terbaik terjadi pada Fold ke-6 sebesar 98.03%. Kemudian rata-rata dari performa model adalah *Accuracy* 93.16%, *Precision* 90.94%, *Recall* 83.96%, dan *F1-Score* 85.58%.
- 3 Selanjutnya dilakukan penelitian untuk mengetahui seberapa baik model yang telah dibuat menurut aspek-aspek yang sudah di tentukan, dan didapatkan hasil *Accuracy* terbaik didapatkan dari aspek Kenyamanan yang bernilai 95.15%, Nilai *Precision* terbaik terdapat pada aspek Kenyamanan dengan nilai 90%, Nilai *Recall* terbaik terdapat pada aspek Kebersihan dengan nilai 85%%, Nilai *F1-Score* terbaik terdapat pada aspek Kenyamanan dengan nilai sebesar 80.81%.

5.2 Saran

Berdasarkan penelitian yang sudah dilakukan, penelitian ini dapat dikembangkan untuk keperluan penelitian selanjutnya dengan saran-saran sebagai berikut:

1. Dalam penelitian ini klasifikasi sentimen hanya dilakukan dengan dua kelas yaitu positif dan negatif, untuk penelitian selanjutnya diharapkan mampu menambahkan kelas sentimen tambahan seperti sentimen netral.
2. Penelitian ini menggunakan 3 klasifikasi Aspek yaitu kebersihan, kelengkapan, dan kenyamanan, untuk penelitian selanjutnya diharapkan menggunakan Aspek yang lebih spesifik, contohnya seperti “ruang ganti”, “kamar, dan “mobil” sebagai sub-kelas yang lebih spesifik dari aspek “fasilitas”.
3. Penelitian menggunakan *Support Vector Machine* sebagai algoritma sebagai model machine learning. Diharapkan selanjutnya dapat menggunakan algoritma yang berbeda atau algoritma *deep learning* seperti BERT, CNN, dan sebagainya.
4. Penelitian menggunakan *Support Vector Machine* dengan *linear kernel*. Untuk penelitian selanjutnya dapat menggunakan kernel yang berbeda untuk mengetahui hasil yang lebih bervariasi seperti kernel RBF, polynomial, dan sigmoid.

DAFTAR PUSTAKA

- Aldisa, R. T., & Maulana, P. (2022). Analisis Sentimen Opini Masyarakat terhadap Vaksinasi Booster COVID-19 dengan perbandingan metode Naïve Bayes, Decision Tree, dan SVM. *Jurnal BITS*, 106 - 109.
- Apriyani, H., & Kurniati. (2020). Perbandingan Metode Naïve Bayes Dan Support Vector Machine Dalam Klasifikasi Penyakit Diabetes Melitus. *Journal of Information Technology Ampera*, 133-143.
- Arum, D. S., Butsianto, S., & Astuti, R. (2023). ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP SEA GAMES 2023 DI TWITTER DENGAN METODE NAÏVE BAYES. *Journal of Information System, Applied, Management, Accounting and Research (JISAMAR)*, 728-738.
- Azis, H., Purnawansyah, Fattah, F., & Putri, I. P. (2020). Performa Klasifikasi K-NN dan Cross-validation pada Data Pasien Pengidap Penyakit Jantung . *ILKOM Jurnal Ilmiah*, 81-86.
- Balakrishnan, V., & Yemoh, E. L. (2014). *Stemming and lemmatization: A comparison of retrieval performances*. Korea: SCEI Seoul Conferences.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python*. Sebastopol: O'Reilly Media.
- Emma, H., Liu, X., & Shi, Y. (2013). The Role of Text Pre-processing in Sentiment Analysis. *Procedia Computer Science*, 26-32.
- Giffari, M. R. (2022). *ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN APLIKASI TANGERANG LIVE MENGGUNAKAN LATENT DIRICHLET ALLOCATION DAN NAIVE BAYES*. Jakarta: Universitas Islam Negeri Syarif Hidayatullah Jakarta.
- Google. (2017, 11 16). *Google Colaboratory*. Retrieved 12 27, 2023, from Google Research: <https://research.google.com/colaboratory/intl/id/faq.html>
- Hafiz, Y. A., & Sudarmilah, E. (2023). IMPLEMENTASI WEB SCRAPING PADA PORTAL BERITA ONLINE. *INISIASI*, 55-60.
- Ibrahim, S. (2023, februari 17). *ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP KEBIJAKAN PEMERINTAH DALAM MENANGANI COVID19 MENGGUNAKAN*

METODE SUPPORT VECTOR MACHINE DAN NAIVE BAYES DENGAN LAPLACE ESTIMATOR. Retrieved from Repository Universitas Negeri Jakarta: <http://repository.unj.ac.id/36857/>

Ipmawati, J., Kusriani, & Luthfi, E. T. (2017). Komparasi Teknik Klasifikasi Teks Mining Pada Analisis Sentimen. *Indonesian journal of Network and Security*, 28-36.

Iskandar, J. W., & Nataliani, Y. (2021). Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek. *Jurnal Resti*, 1120-1126.

Kompas. (2021, 10 19). *Apa itu SEA Games?* Retrieved 9 2, 2023, from Kompas.com: <https://www.kompas.com/skola/read/2021/10/19/100000369/apa-itu-sea-games>

Ladayya, F., Siregar, D., Pranoto, W. E., & Muchtar, H. D. (2022). Analisis Sentimen pada Program Transportasi Publikasi Jaklingko dengan Metode Support Vector Machine. *Jurnal Statistika dan Aplikasinya*, 381-392.

Liu, N., Shen, B., Zhenjiang, Z., Zhang, Z., & Mi, K. (2019). Attention-based Sentiment Reasoner for aspect-based sentiment analysis. *Human-Centered Computing and Information Sciences*, 35-52.

Maulana, I., Apriandari, W., & Pambudi, A. (2023). Analisis Sentimen berbasis Aspek terhadap ulasan aplikasi MyPertamina menggunakan support vector machine. *Jurnal IDEALIS*, 172 - 181.

Nugraha, F. A., Harani, N. H., & Habibi, R. (2020). *Analisis Sentimen Terhadap Pembatasan Sosial Menggunakan Deep Learning*. Bandung: Kreatif Industri Nusantara .

Octaviani, P. A., Wilandari, Y., & Ispriyanti, D. (2014). Penerapan Metode Klasifikasi Support Vector Machine (SVM) Pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang. *Jurnal Gaussian*, 811-820.

Pamungkas, T. J., & Romadhony, A. (2021). Analisis Sentimen Berbasis Aspek Terhadap Ulasan Restoran Berbahasa Indonesia. *e-Proceeding of Engineering*, 4102-4114.

Pang, B., & Lee, L. (2008). OPINION MINING AND SENTIMENT ANALYSIS. *Foundations and Trends in Information Retrieval*, 1-135.

Rahmayanti, N. (2022). *ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN FINANCE TECHNOLOGY MENGGUNAKAN METODE SUPPORT VECTOR MACHINE*. Bandung: UNIKOM.

Septian, J. A., Fahrudin, T., & Nugroho, A. (2019). Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor. *JOURNAL OF INTELLIGENT SYSTEMS AND COMPUTATION*, 43-49.

Subowo, E., Artanto, F. A., Putri, I., & Umaedi, W. (2022). BLTS Muntuk analisis sentimen berbasis aspek pada aplikasi belanja online dengan cicilan. *Jurnal Fasilkom*, 132-140.

Utami, M. A., Silvianti, P., & Masjkur, M. (2023). Algoritme Support Vector Machine untuk Analisis Sentimen Berbasis Aspek Ulasan Game Online Mobile Legends: Bang-Bang. *jurnal Xplore: Journal of Statistics* , 63 - 77.

Utami, S. N. (2021, Oktober 19). *Apa itu SEA Games*. Retrieved from Kompas.com: <https://www.kompas.com/skola/read/2021/10/19/100000369/apa-itu-sea-games>

Yutika, C. H., Adiwijaya, & Faraby, S. A. (2021). Analisis Sentimen Berbasis Aspek pada Review Female Daily. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 422-430.

LAMPIRAN

Lampiran 1. Lembar Persetujuan Dosen



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS NEGERI JAKARTA
FAKULTAS TEKNIK
PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER
Gedung L1 Kampus A Universitas Negeri Jakarta, Jalan Rawamangun Muka, Jakarta 13220
Telepon : 021 – 4712137 Laman: <http://ft.unj.ac.id/ptik/>
Fb: Ptik Ft-unj email: adm.s1ptik@gmail.com

LEMBAR PERSETUJUAN DOSEN PEMBIMBING
SKRIPSI/KOMPREHENSIF/KARYA INOVATIF

Dengan ini kami menyatakan bahwa draft skripsi/komprehensif/karya inovatif dengan judul: **ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP FASILITAS SEA GAMES 2023 MELALUI MEDIA TWITTER MENGGUNAKAN METODE ALGORITMA SUPPORT VECTOR MACHINE.**

Mahasiswa berikut ini :

Nama : Azka Luthfan Rudiana
No. Registrasi : 1512619093
Program Studi : Pendidikan Teknik Informatika dan Komputer
Fakultas : Teknik

Dinyatakan layak dan disetujui untuk diuji pada sidang ujian skripsi/komprehensif/karya inovatif.

Pembimbing I,

Dr. Widodo, S.Kom, M.Kom.
NIP. 197203252005011002

Pembimbing II,

Ressa Dwitias Sari, S.T, M.T.I.
NIP. 198909152019032021

Lampiran 2. Surat Pernyataan Dosen Pembimbing 1



KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI
UNIVERSITAS NEGERI JAKARTA
FAKULTAS TEKNIK

Gedung L Kampus A Universitas Negeri Jakarta, Jalan Rawamangun Muka, Jakarta 13220
Telepon : (62-21) 4751523, 47864808 Fax. 47864808
Laman: <http://ft.unj.ac.id> email: ft@unj.ac.id

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Dr. Widodo, S.Kom, M.Kom.

NIP : 197203252005011002

Adalah Dosen Pembimbing dari mahasiswa:

Nama : Azka Luthfan Rudiana

No.Reg : 1512619093

Program Studi : Pendidikan Teknik Informatika dan Komputer

Angkatan : 2019

Dengan ini saya menyetujui agar mahasiswa dapat mengajukan pendaftaran (~~Seminar-Proposal~~ Skripsi/ Sidang-TA/ Sidang Skripsi/ Sidang-Tesis/ Sidang-PKL)* semester Genap 120, dengan judul "ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP FASILITAS SEA GAMES 2023 MELALUI MEDIA TWITTER MENGGUNAKAN METODE ALGORITMA SUPPORT VECTOR MACHINE".

Demikian surat pernyataan ini saya buat, untuk dipergunakan sebaik-baiknya.

Jakarta, Juli 2024
Dosen Pembimbing

Dr. Widodo, S.Kom, M.Kom.
NIP. 197203252005011002

(*pilih salah satu)

Lampiran 3. Surat Pernyataan Dosen Pembimbing 2



KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI
UNIVERSITAS NEGERI JAKARTA
FAKULTAS TEKNIK

Gedung L Kampus A Universitas Negeri Jakarta, Jalan Rawamangun Muka, Jakarta 13220
Telepon : (62-21) 4751523, 47864808 Fax. 47864808
Laman: <http://ft.unj.ac.id> email: ft@unj.ac.id

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Ressy Dwitias Sari, S.T, M.T.I.
NIP : 198909152019032021

Adalah Dosen Pembimbing dari mahasiswa:

Nama : Azka Luthfan Rudiana
No.Reg : 1512619093
Program Studi : Pendidikan Teknik Informatika dan Komputer
Angkatan : 2019

Dengan ini saya menyetujui agar mahasiswa dapat mengajukan pendaftaran (~~Seminar Proposal Skripsi/ Sidang TA/ Sidang Skripsi/ Sidang Tesis/ Sidang PKI~~)* semester Genap 120, dengan judul "ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP FASILITAS SEA GAMES 2023 MELALUI MEDIA TWITTER MENGGUNAKAN METODE ALGORITMA SUPPORT VECTOR MACHINE".

Demikian surat pernyataan ini saya buat, untuk dipergunakan sebaik-baiknya.

Jakarta, Juli 2024
Dosen Pembimbing II

Ressy Dwitias Sari, S.T, M.T.I.
NIP. 198909152019032021

(*pilih salah satu

Lampiran 4. Lembar Konsultasi Dosen Pembimbing 1



Menerudakan dan
Mewastuhatkan Bangsa

KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET,
DAN TEKNOLOGI
UNIVERSITAS NEGERI JAKARTA
FAKULTAS TEKNIK

PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER

Gedung L1 Kampus A Universitas Negeri Jakarta, Jalan Rawamangun Muka, Jakarta 13220
Telepon : 021 - 47868050 Laman: <http://tl.unj.ac.id/ptik/> Fb: Ptik Ft-unj

LEMBAR KONSULTASI SKRIPSI/KOMPREHENSIF/KARVA INOVATIF

Nama : **Azka Luthfan Rudiana**
No. Registrasi : 1512619093
Program Studi : Pendidikan Teknik Informatika dan Komputer
Judul/Tema : ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP FASILITAS SEA GAMES 2023 MELALUI MEDIA TWITTER MENGGUNAKAN METODE ALGORITMA SUPPORT VECTOR MACHINE

Dosen Pembimbing : **Dr. Widodo, S.Kom, M.Kom**
Tanggal Pertemuan Pertama * : 20 Februari 2023 Paraf KPSJ * :

PERTEMUAN/ TANGGAL	MATERI BAHASAN	PARAF DOSEN	KET.
20 Februari 2023	Pengajuan Judul Skripsi	<i>al</i>	Offline
24 Februari 2023	Bimbingan bab 1	<i>al</i>	Offline
15 Maret 2023	Bimbingan bab 1 dan saran untuk memilih dosen pembimbing 2	<i>al</i>	Offline
19 Juni 2023	Bimbingan bab 2 dan 3	<i>al</i>	Offline
27 Desember 2023	Bimbingan bab 3 revisi	<i>al</i>	Offline
22 Maret 2024	Persetujuan Seminar Proposal	<i>al</i>	Offline
3 April 2024	Bimbingan setelah Seminar Proposal untuk revisi dari dosen penguji	<i>al</i>	Offline
28 Juni 2024	Bimbingan bab 4 dan 5	<i>al</i>	Offline
1 Juli 2024	Persetujuan Sidang Skripsi	<i>al</i>	Offline

Mengetahui,
Koordinator Penyelesaian Studi Jurusan

(Signature)
M. Ficky Duskarnaen, M. Sc
NIP. 197309242006041001

* Diisi dan diparaf paling lambat 2 minggu setelah mendapatkan dosen pembimbing

Lampiran 5. Lembar Konsultasi Dosen Pembimbing 2



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET,
DAN TEKNOLOGI
UNIVERSITAS NEGERI JAKARTA
FAKULTAS TEKNIK
PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER
Gedung L1 Kampus A Universitas Negeri Jakarta, Jalan Rawamangun Muka, Jakarta 13220
Telepon : 021 - 47868050 Laman: <http://ft.unj.ac.id/ptik/> Fb: Ptik Ft-unj

LEMBAR KONSULTASI SKRIPSI//KOMPREHENSIF//KARYA INOVATIF

Nama : **Azka Luthfan Rudiana**
No. Registrasi : **1512619093**
Program Studi : **Pendidikan Teknik Informatika dan Komputer**
Judul/Tema : **ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP FASILITAS SEA GAMES 2023 MELALUI MEDIA TWITTER MENGGUNAKAN METODE ALGORITMA SUPPORT VECTOR MACHINE**

Dosen Pembimbing : **Ressy Dwitias Sari, S.T, M.T.I.**
Tanggal Pertemuan Pertama * : **19 Juni 2023** Paraf KPSJ * :

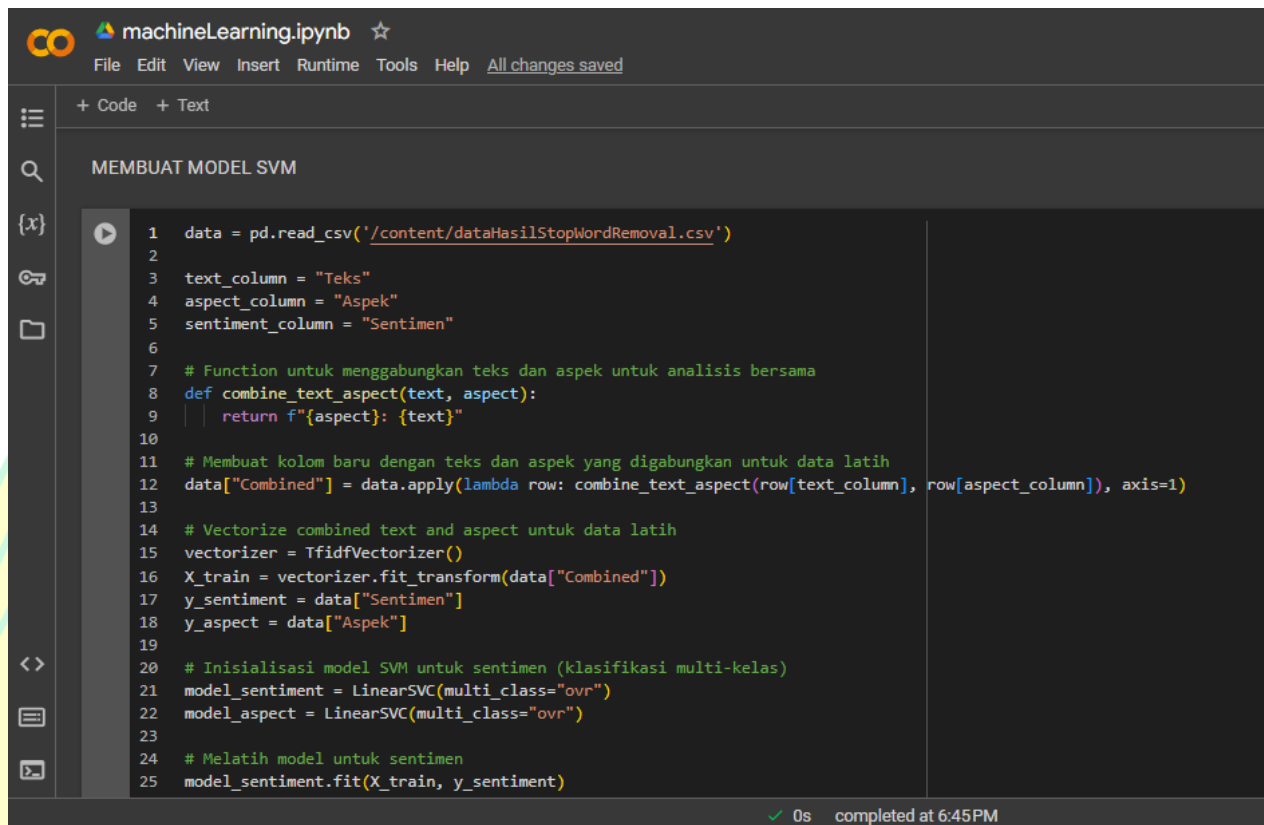
PERTEMUAN/ TANGGAL	MATERI BAHASAN	PARAF DOSEN	KET.
19 Juni 2023	Pengajuan Judul Skripsi		Offline
19 Januari 2024	Bimbingan bab 1		Offline
29 Januari 2024	Bimbingan bab 1 dan 2		Offline
18 Februari 2024	Bimbingan bab 2 dan 3		Offline
22 Maret 2024	Persetujuan Seminar Proposal		Offline
14 Mey 2024	Bimbingan bab 4		Offline
29 Juni 2024	Bimbingan bab 4 dan 5		Online
30 Juni 2024	Persetujuan Sidang Skripsi		Offline

Mengetahui,
Koordinator Penyelesaian Studi Jurusan

M. Ficky Duskarnaen, M. Sc.
NIP. 197309242006041001

* Diisi dan diparaf paling lambat 2 minggu setelah mendapatkan dosen pembimbing

Lampiran 6. Screenshot Kode Pembuatan Model



```
machineLearning.ipynb ☆
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

MEMBUAT MODEL SVM

1 data = pd.read_csv('/content/dataHasilStopWordRemoval.csv')
2
3 text_column = "Teks"
4 aspect_column = "Aspek"
5 sentiment_column = "Sentimen"
6
7 # Function untuk menggabungkan teks dan aspek untuk analisis bersama
8 def combine_text_aspect(text, aspect):
9     return f"{aspect}: {text}"
10
11 # Membuat kolom baru dengan teks dan aspek yang digabungkan untuk data latih
12 data["Combined"] = data.apply(lambda row: combine_text_aspect(row[text_column], row[aspect_column]), axis=1)
13
14 # Vectorize combined text and aspect untuk data latih
15 vectorizer = TfidfVectorizer()
16 X_train = vectorizer.fit_transform(data["Combined"])
17 y_sentiment = data["Sentimen"]
18 y_aspect = data["Aspek"]
19
20 # Inisialisasi model SVM untuk sentimen (klasifikasi multi-kelas)
21 model_sentiment = LinearSVC(multi_class="ovr")
22 model_aspect = LinearSVC(multi_class="ovr")
23
24 # Melatih model untuk sentimen
25 model_sentiment.fit(X_train, y_sentiment)
```

✓ 0s completed at 6:45PM



1. DATA RAW SEBELUM DI *FILTERING*

	A	B	C	D	E	F	G	H	I	J
1976	need_to_r	Wed Nov 1	84	Ah keur turun	hargana oge	RT @Sunda_Bocor:	Gemana hari ini	perolehan emas		
1977	fadhielm	Wed Nov 1		geus loba teuing	komo di pulau	jawa mah RT @Sunda_Bocor	Gmana hari ini	perol		
1978	Lubhee	Tue Nov 1:	Wkwk :D	RT @Sunda_Bocor:	cing pang	bejakeun carana	mareuman capslock	RT @		
1979	efaiti	Tue Nov 1:	RT @Sunda_Bocor:	cing pang	bejakeun carana	mareuman capslock	RT @MALAYSI			
1980	faaadhid	Tue Nov 1:	RT @Sunda_Bocor:	cing pang	bejakeun carana	mareuman capslock	RT @MALAYSI			
1981	jimmytani	Tue Nov 1:	Lg lht	sea games	dgn gor bocor...					
1982	jorgianaa	Tue Nov 1:	Kemaren	gua mau nonton	Sea Games	Basket nya di	Sportmall	deket	tinggal	ngesot
1983	Vicky_Ch	Mon Nov 1	@Erickodoank	wah gawat	malu deh	masak untuk	perhelatan	SEA Games	sampe a	
1984	dianaebel	Mon Nov 1	Yah elu	dodol RT @rivansdq:	@diaanoow	yah gue	gk tau	pala	gue	abis bocor
1985	rivansdq	Mon Nov 1	@diaanoow	yah gue	gk tau	pala	gue	abis	bocor	trus
1986	Estinnads	Mon Nov 1	@sputar	Palembang	lg asik	nntn	sea games?	Wajib	pk	pembalut
1987	obet1996	Sun Nov 1:	Atap	Istora	Bocor	Pertandingan	Terganggu:	13/11/2011	23:00	(SEA Games 2011)
1988	Catatan_S	Sun Nov 1:	Atap	Istora	Bocor	Pertandingan	Terganggu:	13/11/2011	23:00	(SEA Games 2011)
1989	Catatan_S	Sun Nov 1:	Atap	Istora	Bocor	Pertandingan	Terganggu:	13/11/2011	23:00	(SEA Games 2011)
1990	AbidNS95	Sun Nov 1:	Atap	Istora	Bocor	Pertandingan	Terganggu:	13/11/2011	23:00	(SEA Games 2011)
1991	ozy_ev	Sun Nov 1:	Pak	beye	bocor	pak @cheezy_dee:	Atap	Istora	Bocor	Pertandingan
1992	iqbal_hidk	Sun Nov 1:	#kitabisa	RT @rizkisamasajaa:	Atap	Istora	Bocor	Pertandingan	Terganggu:	13/11/2
1993	HENDRRR!	Sun Nov 1:	Atap	Istora	Bocor	Pertandingan	Terganggu:	13/11/2011	23:00	(SEA Games 2011)
1994	dimasrynt	Sun Nov 1:	Atap	Istora	Bocor	Pertandingan	Terganggu:	13/11/2011	23:00	(SEA Games 2011)
1995										
1996										
1997										
1998										
1999										



TENTANG PENULIS



Penulis bernama Azka Luthfan Rudiana, lahir di Sukabumi pada tanggal 27 Januari 2001. Merupakan anak pertama dari 3 bersaudar dari Bapak Moh. Agus Rudiana dan Eva Latiha. Penulis menempuh pendidikan awal di SDN 5 CBH 5, SMPN 2 Kota Sukabumi, SMAN 3 Sukabumi. Pada tahun 2019 Penulis diterima di Universitas Negeri Jakarta, Fakultas Teknik dengan jurusan S1 Pendidikan Teknik Informatika dan Komputer melalui jalur PENMABA.

Kegiatan organisasi yang pernah diikuti oleh penulis selama berkuliah di Universitas Negeri Jakarta adalah KPU UNJ pada tahun 2019. Kemudian BEMP pada tahun 2020/2021 sebagai anggota departemen PSDM.

Kegiatan lainnya yang dilakukan Penulis diluar kegiatan kampus adalah dengan mengikuti program Beasiswa dari BAZNAS (BAZIZ) Pada tahun 2020/2021, mengikuti Praktik Kerja Lapangan di PT. Abdi Teknologi Informasi pada tahun 2021/2022, Mengikuti praktik mengajar di SMKN 24 Jakarta, mengikuti kegiatan kepemiluan sebagai pantarlih dan KPPS untuk pemilu 2024 dan pilkada 2024.

Dalam penyelesaian studinya di Universitas Negeri Jakarta, peneliti melakukan penelitian dalam bentuk Skripsi yang berjudul “Analisis Sentimen Berbasis Aspek Terhadap Fasilitas Sea Games 2023 Melalui Media Twitter Menggunakan Algoritma Support Vector Machine” sebagai salah satu syarat dalam memperoleh gelar sarjana. Untuk menghubungi penulis, dapat menghubungi via azkaluthfan@gmail.com.