

ANALISIS SENTIMEN AKUN TWITTER APEX LEGENDS MENGUNAKAN VADER

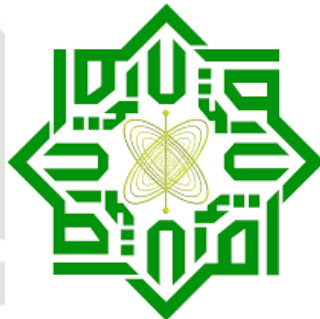
TUGAS AKHIR

Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh

DICKY ABIMANYU

11551102877



UIN SUSKA RIAU

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SULTAN SYARIF QASIM RIAU

PEKANBARU

2022

Hak Cipta Dilindungi Undang-Undang

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

LEMBAR PERSETUJUAN

**ANALISIS SENTIMEN AKUN TWITTER APEX LEGENDS
MENGUNAKAN VADER**

TUGAS AKHIR

Oleh

DICKY ABIMANYU
11551102877

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 11 Juli 2022

Pembimbing 1,



ELVIA BUDIANITA, S.T., M.Cs.
NIP. 19801018200710 1 002

Pembimbing 2,



EKA PANDU CYNTHIA, S.T., M.Kom.
NIP. 19890814 202012 2 012

Hak Cipta Dilindungi Undang-Undang

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

LEMBAR PENGESAHAN

ANALISIS SENTIMEN AKUN TWITTER APEX LEGENDS MENGUNAKAN VADER

TUGAS AKHIR

Oleh:

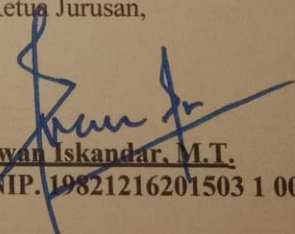
DICKY ABIMANYU
11551102877



Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar sarjana Teknik Informatika
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau
di Pekanbaru, pada tanggal 11 Juli 2022

Pekanbaru, 11 Juli 2022

Mengesahkan,

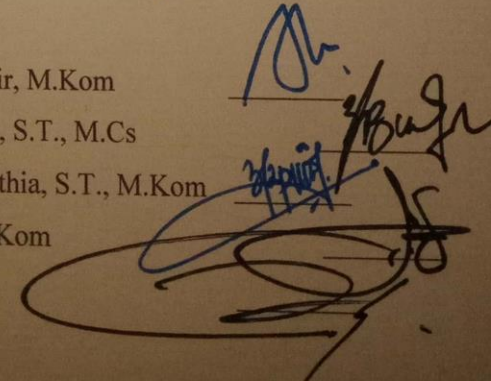
Ketua Jurusan,


Iwan Iskandar, M.T.
NIP. 19821216201503 1 003


Dekan

Dr. Hartono, M.Pd.
NIP. 19640301 1992 03 1 003

DEWAN PENGUJI

Ketua	: DR. Alwis Nazir, M.Kom
Sekretaris	: Elvia Budianita, S.T., M.Cs
Pembimbing II	: Eka Pandu Cynthia, S.T., M.Kom
Penguji I	: Febi Yanto, M.Kom
Penguji II	: Yusra, MT



LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

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

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

Hak Cipta Dilindungi Undang-Undang

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

LEMBAR PERNYATAAN

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

Pekanbaru, 11 Juli 2022

Yang membuat pernyataan,

DICKY ABIMANYU

NIM. 11551102877

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

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

LEMBAR PERSEMBAHAN



Alhamdulillah...

Rasa syukur kuhaturkan kepada-Mu Yaa Allah yang Maha Ber-Ilmu, hanya karena karuniaMu sajalah hamba-Mu akhirnya dapat menyelesaikan Tugas Akhir ini

Tugas Akhir ini saya persembahkan untuk Ayahanda dan Ibunda tercinta.

Tentulah tidak akan tergantikan semua jasa, pengorbanan, tetes keringat dan rasa letih itu, hanya dengan karya kecil ini... Namun semoga dengan ini, saya bisa mengukir sebaris senyum bahagia di hati Ayahanda dan Ibunda tercinta...

Terimakasih untuk semua rangkaian do'a, kasih sayang serta ilmu yang berharga...

Dan tidak lupa saya persembahkan untuk kedua adik tersayang, terimakasih untuk semua dukungan dan doa yang telah diberikan selama ini...

Juga, saya persembahkan untuk semua kerabat, sanak saudara yang mendukung sampai saya bisa berada di titik ini... Semua kesulitan seolah lenyap saat mengingat bahwa aku memiliki dukungan dari kalian semua.

Dan yang terakhir Tugas Akhir ini saya persembahkan kepada Teman, Sahabat, dan semua yang telah mendukung dan membantu saya hingga mencapai titik finish ini...

Hak Cipta Dilindungi Undang-Undang

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



SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini :

Nama : Dicky Abimanyu
NIM : 11551102877
Tempat/Tgl. Lahir : Beringin Jaya, 16 Oktober 1996
Fakultas : Fakultas Sains Dan Teknologi

Judul Skripsi Berbasis Jurnal :

Analisis Sentimen Akun Twitter Apex Legends Menggunakan VADER

Menyatakan dengan sebenar-benarnya bahwa :

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

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

Pekanbaru, 18 Juli 2022

Dicky Abimanyu



NIM : 11551102877

Hak Cipta Dilindungi Undang-Undang

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

Analisis Sentimen Akun Twitter Apex Legends Menggunakan VADER

Dicky Abimanyu¹, Elvia Budianita², Eka Pandu Cynthia³, Febi Yanto⁴, Yusra⁵

Program Studi Teknik Informatika, Fakultas Sains Dan Teknologi^{1,2,3,4,5},

Universitas Islam Negeri Sultan Syarif Qasim Riau

Jl. H.R. Soebrantas No. 155 Km 15, Simpang Baru, Tampan, Pekanbaru 28293

Email : dicky.abimanyu@students.uin-suska.ac.id, elvia.budianita@uin-suska.ac.id,
eka.pandu.cynthia@uin-suska.com, febiyanto@uin-suska.ac.id, yusra@uin-suska.ac.id

Abstrak - Pesatnya peningkatan jasa internet saat ini, ada banyak informasi yang dihasilkan dalam jumlah besar secara terus menerus dalam waktu yang singkat. Akhir-akhir ini, analisis sentimen dengan menggunakan ulasan dan pesan telah menjadi topik penelitian yang populer dibicarakan di bidang *Natural Language Processing*. Selama bertahun-tahun, permainan online telah menjadi suatu aktivitas yang tidak bisa dipisahkan dari sebagian besar orang. Apex Legends adalah salah satu contoh game yang sangat populer di seluruh dunia. Untuk mendapatkan informasi bagaimana pendapat para pemain tentang permainan ini diperlukan analisis sentimen. Pada penelitian ini dilakukan analisis sentimen menggunakan bantuan aplikasi Orange Data Mining dengan metode VADER pada akun twitter Apex Legends menggunakan data sebanyak 500 tweet. Pengujian data dilakukan dengan membandingkan hasil yang didapat menggunakan metode VADER dengan hasil pengujian pakar, yaitu *native speaker* dari Canada dan Amerika. VADER mengklasifikasikan data yang didapatkan melalui twitter berdasarkan nilai compound yang didapat. Penelitian ini menghasilkan kesimpulan yaitu perbandingan dari pengujian menggunakan VADER dan pengujian pakar tidak berbeda jauh, yang mana total persentase dari penggunaan metode VADER untuk menganalisis sentiment dari twitter ini adalah : Positif = 18%, Negatif = 4,8%, Netral = 77,2%. Sedangkan hasil pengujian pakar adalah : Positif = 27%, Negatif = 10,8%, Netral = 62,2%.

Kata kunci: VADER, Apex Legends, Game, Twitter, Uji Pakar

Abstract - With the rapid increase in internet services today, there is a lot of information produced in large quantities continuously in a short time. Recently, sentiment analysis using reviews and messages has become a popular research topic discussed in the *Natural Language Processing* field. Over the years, online gaming has become an activity that cannot be separated from most of the people. Apex Legends is one example of a game that is very popular around the world. To get information on how the players think about the game, sentiment analysis is needed. In this study, sentiment analysis was carried out using the Orange Data Mining application with the VADER method on the Apex Legends twitter account using 500 tweets (data). Data testing is done by comparing the results obtained using the VADER method with the results of expert testing, *native speaker* from Canada and America. VADER classifies the data obtained through twitter based on the compound value obtained. This study concludes that the comparison of testing using VADER and expert testing is not much different, where the total percentage of using the VADER method to analyze sentiment from Twitter is : Positive = 18%, Negative = 4,8%, Neutral = 77,2%. While the results of expert testing is : Positive = 27%, Negative = 10,8%, Neutral = 62,2%.

Keywords : VADER, Apex Legends, Game, Twitter, Expert Test (Uji Pakar)

1. Pendahuluan

Selama bertahun-tahun, game online menjadi hal yang tidak dapat dipisahkan bagi sebagian besar dari kita. Game online menawarkan aktifitas yang menarik bagi orang-orang di rumah yang mencari interaksi sosial, dan riset menunjukkan bahwa terjadinya pertumbuhan besar dalam waktu bermain game. Fasilitas game online yang dapat dimainkan oleh banyak pengguna secara bersamaan ini menjadi daya tarik yang sangat tinggi untuk orang-orang agar terus bermain game online. Dari banyaknya pemain game online di dunia ini banyak diantaranya yang bermain game bergenre FPS *Battle Royale*.

FPS *Battle Royale* adalah salah satu genre game online yang sedang trend sampai saat ini. FPS adalah singkatan dari *First Person Shooter*, sedangkan *Battle Royale* itu sendiri adalah sebuah genre game dimana dalam satu sesi game nya terdiri dari beberapa tim, dan di satu tim ini terdiri dari beberapa pemain yang melakukan elemen bertahan hidup dan eksplorasi untuk menjadi tim terakhir yang bertahan. *Battle Royale* itu sendiri mencakup banyak player, dimana mereka akan memulai permainan dengan perlengkapan seadanya, dan kemudian mereka mulai mencari peralatan perang seperti senjata, alat penyembuh (*medical kit*), peluru, dan lain-lain. Setelah itu mereka akan mulai mengeliminasi pemain dari tim lain sembari menghindari dari "safe area" yang semakin mengecil, lalu pemenangnya akan ditentukan jika salah satu tim menjadi satu satunya tim yang tersisa di pertandingan tersebut[1]. Game *Battle Royale* yang tengah populer saat ini salah satunya adalah Apex Legends. Game ini sangat diminati oleh para player dari seluruh dunia.

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

Salah satu dampak dari perkembangan Internet yaitu pengguna media sosial yang besar. Data dari <https://techno.kompas.com> dan We Are Social, kira-kira 5 milyar orang di dunia aktif diberbagai media sosial dihitung pada April 2022, yang mana angka ini mengalami peningkatan yang signifikan dari tahun lalu dengan kenaikan sebesar 196 juta pengguna[2]. Salah satu sosial media yang populer saat ini yaitu Twitter yang memiliki jumlah pengguna diseluruh dunia hingga 465,1 juta [3]. Banyak orang senang menggunakan sosial media karena bisa dijadikan sebagai sarana banyak orang untuk menuangkan opininya tentang suatu event yang sedang terjadi baik untuk hal personal maupun hal-hal lainnya. Opini-opini tersebut dapat diolah menjadi suatu informasi dengan suatu cara yaitu analisis sentimen.

Analisis sentimen adalah teknik untuk menganalisis opini, sentimen, penilaian dan emosi terhadap suatu entitas seperti produk, jasa, kejadian atau atribut lainnya. Pemikiran dasar dari teknik analisis sentimen adalah untuk mengelompokkan teks, kalimat, atau dokumen kemudian menentukan teks, kalimat, atau dokumen tersebut termasuk ke dalam sentimen atau opini yang positif, negatif, atau netral. Ada 2 metode klasifikasi yang bisa digunakan dalam menganalisis sentimen, yaitu *lexicon based* dan *machine learning based*. Metode *lexicon based* adalah metode yang tidak memerlukan data training atau data yang telah dilabeli namun sudah tersedia dalam kamus lengkap dengan kepolartisan sentimennya. Metode *lexicon based* menggunakan kamus, kosa kata, atau korpus yang sudah ada (yang telah dibuat oleh para ahli linguistik) dan mencocokkannya dengan data yang ingin diklasifikasikan[4]. Sedangkan metode *machine learning* menggunakan algoritma *machine learning* untuk mengklasifikasikan data yang ingin diklasifikasikan. Terdapat korpus leksikal yang cukup terkenal yaitu VADER. VADER merupakan singkatan dari *Valence Aware Dictionary for Sentiment Reasoning*. VADER menggunakan kamus *lexicon* berbahasa inggris (en). VADER *sentiment polarity detection* digunakan untuk proses sentimen analisis karena proses penentuan sentimen yang mudah bersumber dari *dictionary* yang tersedia sehingga bisa mendekati penilaian manusia [4].

Pada jurnal ini, penulis menggunakan bantuan *software* yang bernama Orange Data Mining dengan metode VADER untuk mengklasifikasikan dataset yang didapatkan melalui twitter. Data didapatkan menggunakan teknik *Data Scraping* melalui twitter API yang bisa didapatkan secara gratis di [website developer.twitter.com/apps](https://developer.twitter.com/apps). Setelah itu data akan di *Pre-processing* menggunakan teknik *transformation, tokenization, normalization*, dan yang terakhir yaitu *filtering*. Setelah data melewati *Pre-processing*, data tadi akan dianalisis sentimennya menggunakan metode VADER. Setelah itu data akan ditampilkan dalam bentuk nilai positif, negatif, netral, dan compound. Nilai compound inilah yang digunakan untuk menentukan apakah sentimen yang didapat bermakna positif, negatif, atau netral.

Penelitian terkait yang dilakukan oleh Fauzan Aziz, Abdurrahman Rahim Thaha dan Nizar Abdan Ma'rif [5]. Pada tahun 2022 tentang Analisis Sentimen Destinasi Wisata Geopark Ciletuh yang menghasilkan kesimpulan ulasan para wisatawan dengan sentimen positif mencapai 77,56%, ulasan dengan sentimen netral sebesar 17,15%, dan ulasan dengan sentimen negatif hanya 5,29%. Hasil klasifikasi sentimen emosi dari semua ulasan menunjukkan sentiment emosi didominasi oleh emosi joy dengan jumlah 61,5% atau 3.210 ulasan dari semua ulasan. Hal ini menunjukkan bahwa diatas 50% para wisatawan yang berkunjung sangat senang dan menikmati objek wisata yang ada di Geopark Ciletuh. Pada penelitian lainnya yang dilakukan oleh Putri Amira Sumitro, Rasiban, Dadang Iskandar Mulyana, Wahyu Saputro [6] pada tahun 2021 tentang Analisis Sentimen terhadap vaksin covid-19 di Indonesia pada twitter menggunakan metode *lexicon based* yang menghasilkan kesimpulan didapatkan sentimen positif sebanyak 20,25%, agak positif sebanyak 23,9%, negatif sebanyak 1,88%, agak negatif sebanyak 9,6% dan netral sebanyak 44,36%. Berdasarkan hasil diatas, dapat disimpulkan bahwa opini masyarakat terhadap vaksin Covid-19 pada Twitter lebih cenderung netral.

Berdasarkan permasalahan tersebut, maka penelitian ini akan menggunakan *software* Orange Data Mining dengan menerapkan metode VADER dalam mengklasifikasikan sentimen positif, negatif, atau netral terhadap akun twitter Apex Legends berdasarkan tweet di sosial media twitter.

2. Tinjauan Pustaka

2.1. Analisis Sentimen

Analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, apakah pendapat yang dikemukakan dalam dokumen bersifat positif, negatif atau netral. Penelitian mengenai analisis sentimen telah berkembang sejak tahun 2003 dan merupakan bagian dari *Text Mining* yang merupakan penelitian komputasi berdasarkan sentimen, *emoticon*, pendapat, komentar dan setiap ekspresi yang diungkapkan oleh teks. Analisis sentimen difokuskan untuk review klasifikasi berdasarkan polaritas. Berdasarkan klasifikasi, analisis sentimen dibagi menjadi dua kelompok utama. Yaitu dokumen klasifikasi ke pendapat atau fakta, atau dikenal sebagai klasifikasi subjektivitas (*subjectivity classification*) dan dokumen klasifikasi ke dalam positif atau negatif, atau dikenal sebagai analisis sentimen. Hal ini adalah proses yang penting untuk menentukan dokumen yang memiliki opini dan dokumen yang menyimpulkan opini bernilai positif, negatif maupun netral [7].

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

2.2. Klasifikasi

Menurut Prasetyo, klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia[8]. Klasifikasi merupakan sebuah proses yang mana menjelaskan atau membedakan konsep atau kelas data untuk menemukan model atau fungsi, yang bertujuan untuk dapat memprediksi kelas dari suatu objek yang hasilnya tidak diketahui [9]. Terdapat 2 metode klasifikasi yang biasa digunakan untuk penelitian, yaitu metode klasifikasi berbasis *lexicon* dan metode klasifikasi berbasis *machine learning*. Metode *lexicon based* menggunakan kamus dan kosakata yang sudah ada (yang telah dibuat oleh para ahli linguistik), lalu mencocokkannya dengan data yang ingin diklasifikasikan. Metode *lexicon based* tidak memerlukan data training atau data yang telah dilabeli namun sudah tersedia dalam kamus lengkap dengan kepolaritasan sentimennya [4]. Sedangkan metode *machine learning* menggunakan algoritma *machine learning* untuk mengklasifikasikan data yang ingin diklasifikasikan. Metode *machine learning* memerlukan data yang sudah dilabeli terlebih dahulu karena data yang didapatkan masih harus dibagi menjadi data latih dan data uji.

2.3 Text Mining

Text Mining merupakan proses semi otomatis penggalian pola (pengetahuan dan informasi yang berguna) dari sejumlah sumber data berukuran besar yang tidak terstruktur berupa teks[10]. Teks yang terdiri dari sekumpulan data yang tidak terstruktur akan diproses dan diubah menjadi representasi numerik yang terstruktur[11].

2.4 Text Pre-processing

Pre-processing text bertujuan untuk menyeleksi data tweet yang akan dibersihkan, lalu data tersebut akan digunakan untuk dianalisa / diklasifikasikan sentimennya. Tahap *Pre-processing* text terdiri dari :

1. Transformation

Pada tahap ini, dilakukan tahapan *cleaning*, yaitu penghapusan tanda baca dan karakter yang tidak penting seperti tanda koma, titik, tanda seru, menghilangkan URL dan HTML, mention (@) dan hashtag (#), serta berbagai simbol simbol yang tidak relevan. Lalu selanjutnya huruf pada data yang didapat akan dirubah menjadi huruf kecil (*lowercase*) semua (*case folding*).

2. Tokenizing

Pada tahap ini, data yang didapat akan dipotong menjadi string input berdasarkan tiap kata – kata penyusunnya.

3. Normalisasi

Pada tahap ini, data yang mengandung kata yang tidak baku akan diubah menjadi kata baku menurut kamus bahasa inggris.

4. Filtering

Pada tahap ini, dilakukan filtering kata dengan tahapan *remove stopword*, *Lexicon*, *Numbers*, dan *RegExp*. *Remove stopword* sendiri adalah sebuah kata umum atau berupa kata ganti dan kata sambung yang biasanya muncul dalam jumlah besar dan dianggap tidak memiliki makna, sehingga tidak akan mempengaruhi arti dari kalimat tersebut. Misalnya kata penghubung seperti “and”, “or”, “but”, “in”, “to”, “from”, “which”, “there”, “with”, “will”, “that”, “no”, “want to”, dan lain-lain. Sedangkan *lexicon* adalah filtering data yang memperkecil token yang ada di dalam sebuah kalimat. *Numbers* adalah filtering token yang mengandung angka. Sedangkan *RegExp* adalah filtering data yang menghapus token yang sama dengan token yang ada di widget *Pre-processing*

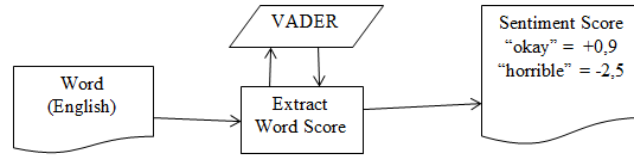
2.5 VADER

Valence Aware Dictionary and Sentiment Reasoner (VADER) adalah metode yang digunakan sebagai model untuk analisis sentimen dan mampu menentukan keragaman data melalui intensitas kekuatan emosional yang ada sesuai dengan kamus data *Lexicon* yang tersedia[12]. VADER diperkenalkan pada tahun 2014 oleh C.J Hutto dan Eric Gilbert yang metode pembentukannya didasarkan pada pendekatan *human-centric*, menggabungkan analisis kualitatif dan validasi empiris menggunakan kebijaksanaan dan penilaian manusia [13].

Kamus *lexicon* dapat digunakan untuk menilai sentimen frasa dan kalimat, tanpa perlu melihat yang lain. Sentimen dapat dikategorikan - seperti {negatif, netral, positif} - atau dapat numerik - seperti kisaran intensitas atau skor. Pendekatan leksikal melihat kategori sentimen atau skor setiap kata dalam sebuah kalimat dan memutuskan kategori atau skor sentimen keseluruhan kalimat itu [14]. Kekuatan dari pendekatan leksikal terletak pada kenyataan bahwa tidak diperlukan melatih model menggunakan data berlabel.

Keuntungan menggunakan VADER *polarity detection* adalah tersedianya kamus yang berisi nilai dari setiap kata. Hasil *Preprocess* text akan di nilai berdasarkan *lexicon* apakah itu positif, negatif atau netral dan menambahkan skor total (*compound*). Beberapa perintah VADER yang menggunakan bahasa pemrograman *python* akan dikerjakan, dan VADER akan memanggil data *lexicon* dari server NLTK untuk menghitung *polarity class sentimen*. Berikut ini adalah *flowchart* proses penentuan polaritas dari suatu kalimat

Gambar 1. Flowchart Penentuan *Polarity Score*



Proses penentuan polaritas kalimat didapatkan dari penyatuan attribute "*compound*" dari setiap kata yang tersedia [15]. Kriteria pengelompokan positif, netral dan negatif yakni jika hasil *compound* lebih dari 0,5 maka dimasukkan kategori positif yang diwakilkan dengan angka 1 lalu jika hasil *compound* terletak diantara -0,5 dan 0,5 maka termasuk kategori netral yang diwakilkan dengan angka 0 dan yang terakhir jika hasil *compound* dibawah -0,5 maka termasuk kategori negatif yang diwakilkan dengan angka -1.

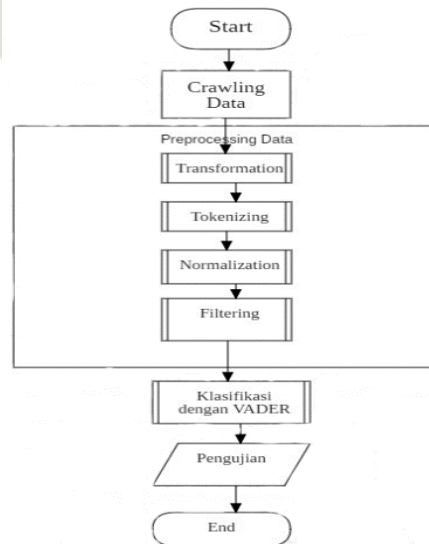
3.6 Orange Data Mining

Orange Data Mining adalah sebuah *Open Source Software* yang digunakan untuk mempermudah proses data mining. Orange dapat digunakan untuk analisis dan visualisasi data eksploratif. Orange mempermudah pemakai bermain dengan data *open source* serta melaksanakan proses *data analytics* secara intuitif [3]. Pada permasalahan riset ini Orange Data Mining menunjukkan sebagian widget untuk mencari data informasi kata yang dominan timbul dari konten status serta pendapat/komentar *account* twitter dengan menciptakan tampilan *word cloud* dari widget Orange Data Mining. Ada beberapa widget yang berguna untuk sentimen analisis, contohnya yaitu widget Sentimen Analisis, widget Twitter untuk mengambil data melalui twitter API, widget *select column* untuk memisahkan data hasil web scraping yang tidak diperlukan seperti nama pengguna, tanggal, dll.

3. Metode Penelitian

Berikut adalah tahapan yang dilakukan dalam jurnal yang berjudul "Analisis Sentimen Akun Twitter Apex Legends Menggunakan VADER" :

Gambar 2. Metode Penelitian



3.1 Crawling Data

Tahapan ini merupakan tahapan pengumpulan data yang diperlukan untuk menganalisa dan memperoleh data-data serta informasi yang berhubungan dengan penelitian yang akan dilakukan. Data yang dikumpulkan berasal dari Tweet pada Twitter akun Apex Legends (@PlayApex). Data yang dikumpulkan dari *server* Twitter menggunakan Twitter API (*Application Programming Interface*) dalam widget Orange Data Mining. Data yang telah dikumpulkan disimpan dalam data corpus, selanjutnya data akan diolah ke tahap preprocessing.

3.2 Text Pre-processing

Setelah data disimpan ke dalam widget data corpus, selanjutnya data akan masuk ke tahap *Pre-processing* text. Langkah - langkah dalam analisis yang diterapkan secara berurutan dapat diaktifkan pada Orange Data Mining dalam widget Preprocess Text. Di bawah ini tahapan dalam preprocess text dilakukan sebelum text

Hak Cipta Dilindungi Undang-Undang
 © Hak cipta milik Universitas Islam Riau
 State Islamic University of Sultan Al-Ma'mun Riau
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang
 © Hak Cipta dan Hak Moral dilindungi Undang-Undang
 © Sultan Muhammad Syarif Qasim Sultan Syarif Qasim
 1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Analisis menggunakan Orange Data Mining :

1. Transformation

Tahapan pertama dari *Text Pre-processing* ini adalah *transformation*, yaitu proses penginputan data dan perubahan data yang didapat sesuai menu pada widget yang tersedia, yaitu merubah semua data menjadi huruf kecil (*lowercase*), menghapus semua kata yang mengandung aksentuasi bahasa (*remove accent*), menghapus kata yang mengandung link (*parse HTML* dan *remove url*)

2. Tokenization

Selanjutnya data akan memasuki tahap tokenization, yaitu tahapan pemisahan kalimat menjadi bentuk per kata (token) dengan memilih menu *Word Punctuation*, dimana menu ini memisahkan kalimat menjadi bentuk per kata dan tetap mempertahankan tanda baca.

3. Normalization

Setelah melakukan tokenization, langkah ketiga dari preprocess text adalah normalization, dalam hal ini *lemmatization text*. Text yang telah dipisahkan kata per kata akan menjadi sebuah text yang berdiri sendiri dalam sebuah kalimat, dan biasanya mengandung penulisan yang kurang sempurna. Dengan proses normalisasi ini text tersebut akan diketahui maknanya dengan menggunakan *WordNet Lemmatizer* yang menerapkan jaringan sinonim kognitif untuk token berdasarkan pada basis data lexicon bahasa Inggris yang besar dari NLTK (Natural Language Toolkit).

4. Filtering

Proses terakhir dari preprocessing text ini adalah tahap filtering, yaitu menghapus atau menyimpan pilihan kata yang ada. Proses pemisahan ini adalah proses dimana kata dan simbol yang tidak digunakan untuk analisis sentimen akan dihapus. Tahapan menu yang ada di Orange Data Mining adalah sebagai berikut :

a. Stopwords

Stopwords menghapus kata penghubung yang terdapat dalam sebuah kata atau kalimat, misalnya menghapus “and”, “or”, “that”, dll. NLTK server telah menyediakan daftar *stopwords* yang dapat diunduh untuk keperluan bahasa. Namun dalam penelitian ini *stopwords* secara default menggunakan bahasa Inggris. Hal ini akan menyaring hanya kata - kata kunci yang disediakan. Berikut ini contoh beberapa *stopwords* yang akan dihilangkan, *stopwords* ini terkoneksi dengan data – data pada server NLTK :

Tabel 1 Daftar English Stopwords
 English Stopwords List

<i>I, me, my, myself, we, our, ours, ourselves, you, he, him, his, himself, she, she's, her, hers, herself, it, it's, its, itself, they, them, their, theirs, themselves, what, which, who, whom, etc.</i>
--

b. RegExp

RegExp menghapus kata-kata yang cocok dengan ekspresi reguler. Secara default, *regexp* diatur untuk menghapus tanda baca. Berikut ini daftar tanda baca yang akan dihapus dari corpus : `\.,!;:|!|\?|\(|\)|\|\\+|'!"|_|\]|!|\'|\\.\|\-|_|_|\$|&|*|#|@|%|_|>|<|\/|\|\\|`. Kita bisa menambahkan sendiri tanda baca yang ingin dihapus jika tidak ada didalam daftar *default regexp*.

3.3 Analisis Sentimen

Data hasil Preprocess text selanjutnya akan dianalisis dengan metode analisis sentimen berbasis rule-based sentiment analysis yaitu VADER. Proses analisis data bertujuan untuk mendapat kategori class sentimen (positif, negatif, netral). Text hasil Preprocessing akan di nilai berdasarkan lexicon apakah itu positif, negatif atau netral dan menambahkan skor total (compound). VADER bekerja dengan menggunakan perintah bahasa pemrograman python, dan yang paling penting adalah ketika VADER memanggil data lexicon dari server NLTK untuk menghitung *polarity class* sentimen. VADER memberikan peringkat text pada skala dari “(-4) Sangat Negatif” hingga “(4) Sangat Positif”, dengan penyisihan “(0) Netral”. Skor sentimen sebuah kalimat dihitung dengan menjumlahkan skor sentimen dari setiap kata yang terdaftar dalam VADER lexicon dalam kalimat. *Score polarity* setiap class akan dinormalkan antara -1 (paling negatif) dan 1 (paling positif) dengan nilai 0 adalah netral. Hasil *score compound* mewakili total skor sentimen yaitu dimana nilai -1 adalah yang paling negatif dan 1 paling positif. Rumus normalisasi score yang digunakan oleh Hutto adalah :

$$\frac{x}{\sqrt{x^2 + \alpha}}$$

dimana x adalah jumlah dari nilai sentimen dari kata-kata penyusun kalimat dan alpha adalah konstanta normalisasi yang set secara default ke 15. Dengan demikian, analisis sentimen VADER berfungsi paling baik pada dokumen pendek, seperti tweet dan kalimat, bukan pada dokumen besar.

3.4 Pengujian



Hak Cipta Dilindungi Undang-Undang

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

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

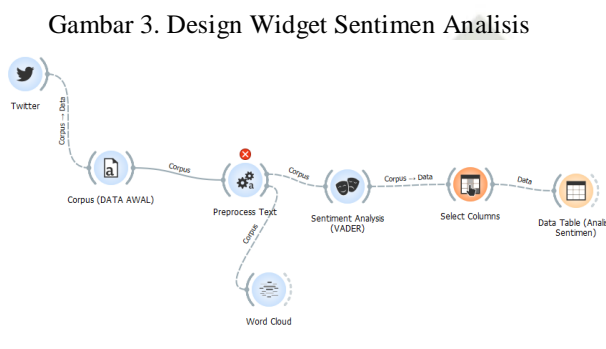
© Himpunan Ilmiah UIN Suska Riau

State Islamic University of Sultanarifki Riau

Penelitian dilakukan dengan membandingkan hasil yang didapat dari metode VADER dengan hasil yang didapat dari Pakar, yaitu seorang Native Speaker yang berasal dari Canada dan America bernama Mr. David vans (Canada) dan Mrs. Mary Stump (America). Hasil analisa pakar didapatkan langsung dengan melakukan pendekatan emosi dan jenis dari kalimat yang didapatkan melalui *twitter*.

Hasil dan Pembahasan
1.1 Skenario Penelitian

Desain *widget* sentimen analisis dengan menggunakan Orange Data Mining digambarkan seperti di bawah ini :



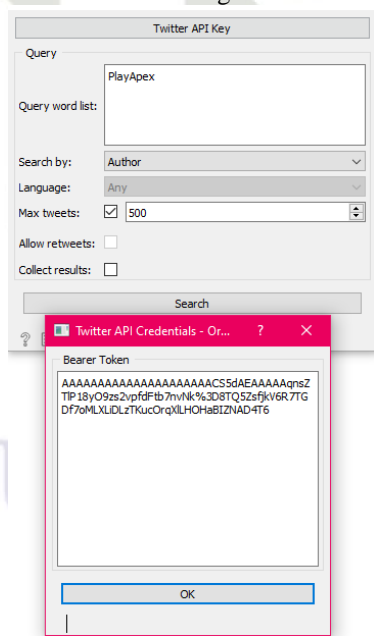
Gambar 3. Design Widget Sentimen Analisis

Data yang dicrawling dari twitter akan disimpan dan dianalisis satu persatu berdasarkan objek. Kemudian data dihubungkan pada widget sesuai dengan apa yang akan digunakan pada penelitian sehingga menghasilkan widget seperti gambar tersebut.

a) Data Crawling

Data yang diambil pada penelitian ini adalah tweet yang berasal dari akun twitter resmi Apex Legends (@PlayApex) di twitter. Data diambil dari 500 tweet dengan menuliskan nama akun twitter "PlayApex" dari API dan *bearer token* twitter. Dalam Orange Data mining data inputan bisa disebut dengan corpus. Widget corpus yaitu kumpulan dokumen yang bisa menyajikan jumlah baris kalimat, serta memastikan mana fitur yang akan dan tidak akan diinput untuk keperluan analisis.

Gambar 4. Widget Twitter



b) Preprocessing text

Sebelum analisis text dilakukan, text diolah terlebih dahulu dengan cara di *preprocessing text*. Langkah – langkah dalam analisis diterapkan secara berurutan serta dinonaktifkan pada Orange Data Mining dalam widget *preprocess text*. Berikut ini tahapan dari preprocess text pada Orange Data Mining sebelum text



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

tersebut dianalisis sentimennya :

1. *Transformation*

Proses penginputan data dan merubah data yang didapat sesuai menu pada widget yang tersedia, yaitu merubah semua data menjadi huruf kecil (*lowercase*), menghapus semua kata yang mengandung aksen bahasa (*remove accent*), menghapus kata yang mengandung link (*parse HTML* dan *remove url*).

2. *Tokenization*

Proses pemisahan kalimat menjadi bentuk per kata (token) dengan memilih menu *Word Punctuation*, dimana menu ini memisahkan kalimat menjadi bentuk per kata dan tetap mempertahankan tanda baca.

3. *Normalization*

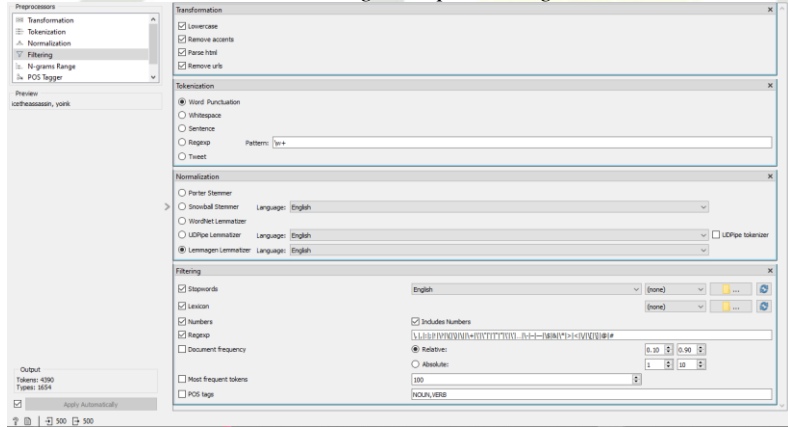
Proses pemisahan kata per kata akan menjadi sebuah text yang berdiri sendiri dalam sebuah kalimat, dan biasanya mengandung penulisan yang kurang sempurna. Dengan proses normalisasi ini text tersebut akan diketahui maknanya dengan menggunakan *WordNet Lemmatizer* yang menerapkan jaringan sinonim kognitif untuk token berdasarkan pada basis data lexicon bahasa Inggris yang besar dari *NLTK* (*Natural Language Toolkit*).

4. *Filtering*

Proses menghapus atau menyimpan pilihan kata yang ada. Proses pemisahan ini adalah proses dimana kata dan simbol yang tidak digunakan untuk analisis sentimen akan dihapus. Tahapan *filtering* yang digunakan pada widget ini adalah :

- a. *Stopwords*, yaitu menghapus kata *stopwords* dari teks (misalnya, menghapus 'and', 'or', 'this' ...). ini dapat dilakukan dengan memilih bahasa yang akan disaring (*filter*). *Stopwords* didapatkan secara default menggunakan bahasa inggris melalui *server NLTK*.
- b. *Lexicon*, yaitu hanya mengambil token yang ada di kamus *lexicon*.
- c. *Numbers*, yaitu menghapus angka yang tidak memiliki arti untuk dianalisis sentimennya.
- d. *RegExp*, yaitu menghapus token yang sesuai dengan ekspresi reguler yang tersedia dalam *Orange Data Mining*.

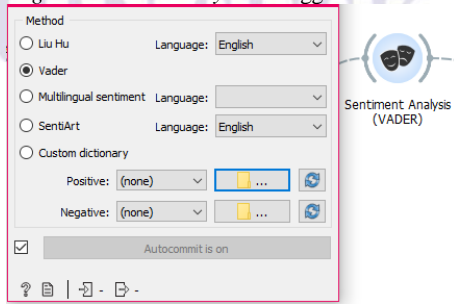
Gambar 5. *Widget Preprocessing Text*



c) Analisis Sentimen

Proses analisis melibatkan algoritma *VADER* untuk mengklasifikasikan *polarity* (*class sentiment*) yaitu positif, negatif, dan neutral dengan skor total yaitu *compound*.

Gambar 6. *Widget Sentiment Analysis Menggunakan VADER*



ji pakar yang merupakan *Native Speaker English* (Mr. David Evans & Ms. Mary Stump), total persentase analisis sentimen dari twitter ini adalah : Positif = 27%, Negatif = 10,8%, Netral = 62,2%. Akurasi pengujian antara metode VADER dengan pengujian pakar adalah 326 dari total 500 data : 65,2%.

Kesimpulan

Dari hasil analisis sentimen yang dilakukan pada akun twitter Apex Legends, didapatkan kesimpulan sebagai berikut :

1. Metode analisis VADER dapat digunakan sebagai metode analisis sentimen yang menghasilkan tiga class sentimen yaitu sentimen positif sebesar 18%, negatif 4,8%, dan neutral sebesar 77,2%.
2. Hasil terbesar berdasarkan analisis sentimen mengenai game Apex Legends adalah netral.
3. Berdasarkan hasil pengujian dengan pakar yaitu Mr. David Evans dan Ms. Mary Stump, klasifikasi menggunakan metode VADER memiliki akurasi sebesar 65,2% dari 500 data uji.

Daftar Pustaka

- [1] "Battle royale game - Wikipedia." .
- [2] "Jumlah Pengguna Internet Dunia Tembus 5 Miliar, Online Hampir 7 Jam Sehari." .
- [3] "The Latest Twitter Statistics: Everything You Need to Know — DataReportal – Global Digital Insights." .
- [4] J. Garay, R. Yap, and M. J. Sabellano, "An analysis on the insights of the anti-vaccine movement from social media posts using k-means clustering algorithm and VADER sentiment analyzer," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 482, no. 1, 2019.
- [5] F. Aziz, A. R. Thaha, and N. Abdan, "Analisis Sentimen Destinasi Wisata Geopark Ciletuh Ujunggenteng menjadi paling terendah dibandingkan kawasan wisata Bandung dan Cirebon . Ketimpangan tingkat," vol. 27, no. 1, pp. 60–68, 2022.
- [6] P. A. Sumitro, Rasiban, D. I. Mulyana, and W. Saputro, "Analisis Sentimen Terhadap Vaksin Covid-19 di Indonesia pada Twitter Menggunakan Metode Lexicon Based," *J-ICOM - J. Inform. dan Teknol. Komput.*, vol. 2, no. 2, pp. 50–56, 2021.
- [7] F. Nurhuda, S. W. Sihwi, and A. Doewes, "Analisis sentimen masyarakat terhadap calon Presiden Indonesia 2014 berdasarkan opini dari Twitter menggunakan metode Naive Bayes Classifier," *ITSmart J. Teknol. dan Inf.*, vol. 2, no. 2, pp. 35–42, 2013.
- [8] E. Prasetyo, "Fuzzy K-Nearest Neighbor in Every Class Untuk Klasifikasi Data Fuzzy K-Nearest Neighbor in Every Class," *Semin. Nas. Tek. Inform. (SANTIKA 2012)*, no. November, pp. 1–5, 2015.
- [9] A. Aradea, I. Supriana, and K. Surendro, "Segmentasi Koleksi Data Untuk Mendukung Performansi Sistem," *J. Ilm. Teknol. Infomasi Terap.*, vol. 2, no. 2, 2016.
- [10] N. R. S. Purba and F. Riandari, "Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Analisis Keranjang Belanja Transaksi Penjualan Pada PT Madu Kembang Joyo," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 4, no. 1, 2021.
- [11] "R: Mining spatial, text, web, and social media data - Biter Makhabel, Pradeepta Mishra, Nathan Hanneman, Richard Heimann - Google Buku." .
- [12] Elbagir and J. Yang, "Language Toolkit and VADER Sentiment," *Proc. Int. MultiConference Eng. Comput. Sci.*, vol. 0958, pp. 12–16, 2019.
- [13] Hutto, C.J. and Gilbert, "VADER: A Parsimonious Rule-based Model for," *Eighth Int. AAI Conf. Weblogs Soc. Media*, p. 18, 2014.
- [14] H. Saif, M. Fernandez, Y. He, and H. Alani, "Evaluation datasets for Twitter sentiment analysis a survey and a new dataset, the STS-Gold," *CEUR Workshop Proc.*, vol. 1096, pp. 9–21, 2013.
- [15] M. Ghiassi and S. Lee, "A domain transferable lexicon set for Twitter sentiment analysis using a supervised machine learning approach," *Expert Syst. Appl.*, vol. 106, pp. 197–216, 2018.

Hak Cipta Dilindungi Undang-undang
 © Hea Kcipia a mlii H a P N s k a R i a u
 © Saia e h a n a n c U n i v e r s i t y o f S u l t a n S y a r i f K a r i m R i a u

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

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