

ANALISIS SENTIMEN *CYBERBULLYING* PADA KOMENTAR INSTAGRAM DENGAN MENGGUNAKAN METODE ALGORITMA *NAI'VE BAYES* DAN *K-NEAREST NEIGHBOR*

TUGAS AKHIR

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

Oleh:

FITRI ANISA NIRMALA

12050322148



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

2024

Hak Cipta Dilindungi Undang-Undang

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

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- 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 CYBERBULLYING PADA KOMENTAR
INSTAGRAM DENGAN MENGGUNAKAN METODE
ALGORITMA NAÏVE BAYES DAN K-NEAREST NEIGHBOR**

TUGAS AKHIR

Oleh:

FITRI ANISA NIRMALA

12050322148

Telah diperiksa dan disetujui sebagai laporan tugas akhir
di Pekanbaru, pada tanggal 06 Juni 2024

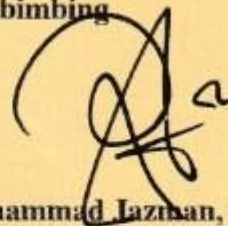
Ketua Program Studi



Eki Saputra, S.Kom., M.Kom.

NIP. 198307162011011008

Pembimbing



Muhammad Lazman, S.Kom., M.InfoSys.

NIP. 198206042015031004

LEMBAR PENGESAHAN

ANALISIS SENTIMEN CYBERBULLYING PADA KOMENTAR INSTAGRAM DENGAN MENGGUNAKAN METODE ALGORITMA NAÏVE BAYES DAN K-NEAREST NEIGHBOR

TUGAS AKHIR

Oleh:

FITRI ANISA NIRMALA

12050322148

Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau
di Pekanbaru, pada tanggal 21 Mei 2024

Pekanbaru, 21 Mei 2024

Mengesahkan,



Ketua Program Studi

Eki Saputra, S.Kom., M.Kom.

NIP. 198307162011011008

DEWAN PENGUJI:

Ketua : Syaifullah, SE., M.Sc.

Sekretaris : Muhammad Jazman, S.Kom., M.InfoSys.

Anggota 1 : Nesdi Evrilyan Rozanda, S.Kom., M.Sc.

Anggota 2 : Febi Nur Salisah, S.Kom., M.Kom.

LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

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

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

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

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

Lampiran Surat :

Nomor : Nomor 25/2021

Tanggal : 10 September 2021

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini:

Nama : Fitri Anisa Nirmala

NIM : 12050322148

Tempat/Tgl. Lahir : Kabun, 16 Februari 2002

Fakultas/Pascasarjana : Sains dan Teknologi

Prodi : Sistem Informasi

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

ANALISIS SENTIMEN CYBERBULLYING PADA KOMENTAR
INSTAGRAM DENGAN MENGGUNAKAN METODE ALGORITMA
NAVE BAYES DAN K-NEAREST NEIGHBOR

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 ~~Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya*~~ saya ini, saya nyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan ~~Disertasi/Thesis/Skripsi/(Karya Ilmiah lainnya)*~~ saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

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

Pekanbaru, 06 Juni 2021

embuat pernyataan



Fitri Anisa Nirmala

NIM : 12050322148

*pilih salah satu sesuai jenis karya tulis

LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Dengan menyebut nama Allah yang maha pengasih lagi maha penyayang

Assalamu 'alaikum Warahmatullahi Wabarakatuh

Alhamdulillah Rabbil 'Alamin, segala puji bagi Allah *Subhanahu Wa Ta'ala*

sebagai bentuk rasa syukur atas segala nikmat yang telah diberikan tanpa ada kekurangan sedikitpun. Shalawat beserta salam tak lupa pula kita ucapkan kepada Nabi Muhammad *Shallallahu 'Alaihi Wa Sallam* dengan mengucapkan "*Allahumma Sholli 'ala Sayyidina Muhammad Wa 'ala Ali Sayyidina Muhammad*". Semoga kita semua selalu senantiasa mendapat syafaat-Nya di dunia maupun di akhirat, aamiin ya rabbal'aalamiin.

Penulis mempersembahkan karya kecil ini sebagai salah satu bentuk bakti, rasa terima kasih, dan hormat kepada Ayah dan Ibu tercinta. Terima kasih yang tak terhingga karena telah merawat dan membesarkan saya dengan setulus hati dan penuh perjuangan hingga saya bisa sampai tahap saat ini. Berkat doa dan kasih sayangmu, anakmu telah berhasil memperoleh gelar sarjana seperti yang engkau harapkan. Tiada apapun di dunia ini yang dapat membalas semua jasa-jasa dan pengorbananmu. Untuk itu saya anakmu ini selalu mendoakan yang terbaik untuk Ayah dan Ibu agar bahagia dunia dan akhirat serta diberikan tempat istimewa di sisi-Nya kelak. Selanjutnya saya juga berterima kasih yang tak terhingga kepada saudara kandung tercinta saya yaitu Abang dan Adik yang telah memberikan saya pelajaran dan pemahaman mengenai indahnya kehidupan yang damai sebagai saudara.

Kepada Bapak dan Ibu Dosen Program Studi Sistem Informasi Universitas Islam Negeri Sultan Syarif Kasim Riau yang telah memberikan ilmu pengetahuan, pengalaman, dan kebaikan selama perkuliahan, saya ucapkan terima kasih banyak dan semoga menjadi amal jariyah. Aamiin.

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

Wassalamu 'alaikum Warahmatullahi Wabarakatuh

KATA PENGANTAR

Alhamdulillah Rabbil 'Alamin, bersyukur kehadirat Allah *Subhanahu Wa Ta'ala* atas segala rahmat dan karunia-Nya sehingga peneliti dapat menyelesaikan Tugas Akhir ini. Shalawat serta salam kita ucapkan kepada Nabi Muhammad *Shalallahu 'Alaihi Wa Sallam* dengan mengucapkan *Allahumma Sholli 'Ala Sayyidina Muhammad Wa'Ala Ali Sayyidina Muhammad*. Tugas Akhir ini dibuat sebagai salah satu syarat untuk mendapatkan gelar Sarjana Komputer di Program Studi Sistem Informasi Universitas Islam Negeri Sultan Syarif Kasim Riau.

Pada penulisan Tugas Akhir ini, terdapat beberapa pihak yang sudah berkontribusi dan mendukung peneliti baik berupa materi, moril, dan motivasi. Peneliti ingin mengucapkan banyak terima kasih kepada:

1. Bapak Dr. H. Hairunas, M.Ag sebagai Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd sebagai Dekan Fakultas Sains dan Teknologi.
3. Bapak Eki Saputra, S.Kom., M.Kom sebagai Ketua Program Studi Sistem Informasi.
4. Bapak Muhammad Jazman, S.Kom., M.InfoSys sebagai dosen pembimbing Tugas Akhir ini.
5. Bapak Syaifullah, SE., M.Sc sebagai Ketua Sidang peneliti yang telah banyak memberikan arahan, masukan, nasihat, serta motivasi dalam penyelesaian Tugas Akhir ini juga dalam perkuliahan.
6. Bapak Nesdi Evrilyan Rozanda, S.Kom., M.Sc sebagai Penguji I peneliti yang telah banyak memberikan arahan, masukan, nasihat, serta motivasi dalam penyelesaian Tugas Akhir ini juga dalam perkuliahan.
7. Ibu Febi Nur Salisah, S.Kom., M.Kom sebagai Penguji II peneliti yang telah banyak memberikan arahan, nasihat, masukan, serta motivasi dalam penyelesaian Tugas Akhir ini juga dalam perkuliahan.
8. Seluruh Bapak dan Ibu Dosen Program Studi Sistem Informasi yang telah banyak memberikan ilmunya kepada peneliti. Semoga ilmu yang diberikan dapat peneliti amalkan dan menjadi amal jariyah.
9. Superhero dan Panutanku, Ayahanda Jumadi terima kasih selalu berjuang untuk kehidupan penulis, mendidik, memotivasi, serta memberikan dukungan hingga penulis mampu menyelesaikan studinya hingga sarjana.
10. Pintu surgaku, Ibunda Nurrohmah yang tidak henti-hentinya memberikan doa, kasih sayang serta sujudnya yang selalu menjadi doa untuk kesuksesan anak-anaknya.



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.

11. Kepada cinta kasih saudara kandung saya, Rudi Irwansyah, S.T., Hendra Gunardi, S.P., dan Ferdi Hasim Alkarim . Terima kasih telah memberikan perhatian, semangat, dukungan, serta doa kepada peneliti.
12. Kepada Sahabat Seperjuangan Terima kasih atas segala bentuk support, canda, tawa dan tangis air mata yang dilalui bersama dalam menempuh pendidikan ini.

Semoga segala doa dan dorongan yang telah diberikan selama ini menjadi amal kebajikan dan mendapat balasan setimpal dari Allah *Subhanahu Wa Ta'ala*. Peneliti menyadari bahwa penulisan Tugas Akhir ini masih banyak terdapat kekurangan dan jauh dari kata sempurna. Untuk itu kritik dan saran yang membangun sangat diharapkan untuk kesempurnaan Tugas Akhir ini dan semoga laporan ini bermanfaat bagi kita semua. Akhir kata peneliti ucapkan terima kasih.

Pekanbaru, 06 Juni 2024
Penulis,

FITRI ANISA NIRMALA
NIM. 12050322148

UIN SUSKA RIAU

8 Mei 2024

: 1997/LoA/JUTIF/VI/2024

: -

: Surat Pemberitahuan Bukti Penerimaan Paper

1. Fitri Anisa Nirmala, Informatics, Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
2. Muhammad Jazman, Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
3. Nesdi Evrilyan Rozanda, Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
4. Febi Nur Salisah, Informatics, Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

Jurnal Teknik Informatika (JUTIF) adalah jurnal nasional terdaftar di LIPI dengan P-ISSN : 2723-3863 dan e-ISSN : 2723-3871 serta terakreditasi SINTA 3 berdasarkan Surat Keputusan Direktur Jenderal Pendidikan Tinggi, Riset, Dan Teknologi Kementerian Pendidikan, Kebudayaan, Riset, Dan Teknologi Republik Indonesia Nomor 164/E/KPT/2021.

Kami mengucapkan terima kasih atas pengiriman artikel ilmiah ke **Jurnal Teknik Informatika (JUTIF)** dengan data artikel sebagai berikut:

Judul

CYBERBULLYING SENTIMENT ANALYSIS OF INSTAGRAM COMMENTS USING NAÏVE BAYES CLASSIFIER AND K-NEAREST NEIGHBOR ALGORITHM METHODS

Penulis


1. Fitri Anisa Nirmala
2. Muhammad Jazman
3. Nesdi Evrilyan Rozanda
4. Febi Nur Salisah

Berdasarkan hasil review, artikel tersebut dinyatakan **DITERIMA** untuk dipublikasikan pada **Jurnal Teknik Informatika (JUTIF) UNIVERSITAS JENDERAL SOEDIRMAN (UNSOED)** Volume 5 Nomor 6 Desember 2024.

Demikian surat pemberitahuan ini, kami mengucapkan terima kasih atas partisipasi dan kerja sama yang baik.

Chief Editor,

 **JUTIF**
Jurnal Teknik Informatika


Dr. Lasmedi Afuan, S.T., M.Cs
NIP. 198505102008121002

CYBERBULLYING SENTIMENT ANALYSIS OF INSTAGRAM COMMENTS USING NAÏVE BAYES CLASSIFIER AND K-NEAREST NEIGHBOR ALGORITHM METHODS

Fitri Anisa Nirmala^{*1}, Muhammad Jazman², Nesdi Evrilyan Rozanda³, Febi Nur Salisah⁴

^{1,2,3,4} Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau
Email: ¹2050322148@student.uin-suska.ac.id, ²jazman@uin-suska.ac.id, ³nesdi.rozanda@uin-suska.ac.id, ⁴febinursalisah@uin-suska.ac.id

(Article received: date; Revision: date; published: date)

Abstract

The high number of social media users presents major threats and risks, such as cyberbullying. Cyberbullying or cyberbullying is one of the negative impacts of the rapid development of technology and social media. Sentiment Analysis is a technique for extracting text data to obtain information about positive, neutral or negative sentiment. One of Indonesian social media that often gets user sentiment through social media is Instagram. By using the Text Mining technique, the classification method will determine whether a sentiment is positive, neutral or negative. This research uses the Naïve Bayes Classifier (NBC) and K-Nearest Neighbor (KNN) methods with tf-idf weighting accompanied by the addition of an emotional icon (emoticon) conversion feature to determine the existing sentiment classes from tweets about Instagram users. The results of calculations using the first three methods using the Partitionong model, the results using the Naive Bayes method, get an accuracy value of 91.25%, a recall value of 92% and a precision value of 90% and calculations using the KNN method have an accuracy value of 67%, a recall value of 49% and a precision value of 34%. So it can be concluded that the Naïve Bayes Classifier algorithm has the best performance.

Keywords: Sentiment Analysis, Cyberbullying K-Nearest Neighbor, Naïve Bayes Classifier

ANALISIS SENTIMEN CYBERBULLYING PADA KOMENTAR INSTAGRAM DENGAN MENGGUNAKAN METODE ALGORITMA NAÏVE BAYES CLASSIFIER DAN K-NEAREST NEIGHBOR (KNN)

Abstrak

Tingginya jumlah pengguna media sosial menghadirkan ancaman dan risiko yang besar, seperti perundungan siber (*cyberbullying*). *Cyberbullying* atau perundungan siber merupakan salah satu dampak negatif dari pesatnya perkembangan teknologi serta sosial media. Analisis Sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negatif. Salah satu media sosial Indonesia yang sering mendapatkan sentimen pengguna melalui media sosial adalah Instagram. Dengan menggunakan teknik Text Mining metode klasifikasi, akan diketahui suatu sentimen bernilai positif atau negatif. Penelitian ini menggunakan metode *Naïve Bayes Classifier* (NBC) dan *K-Nearest Neighbor* (KNN) dengan pembobotan TF-IDF disertai penambahan fitur konversi ikon emosi (*emoticon*) untuk mengetahui kelas sentimen yang ada dari tweet tentang pengguna instagram. Hasil perhitungan dengan menggunakan ketiga metode yang pertama dengan menggunakan model Partitionong hasil dengan metode NBC mendapatkan nilai *accuracy* 91.25%, nilai *recall* 92% dan nilai *precision* 90% dan perhitungan dengan metode KNN nilai *Accuracy* 67%, nilai *recall* 49% dan nilai *precision* 34%. Sehingga dapat disimpulkan Algoritma metode *Naïve Bayes Classifier* merupakan performa terbaik.

Kata kunci: Analisis Sentimen, Cyberbullying K-Nearest Neighbor, Naïve Bayes Classifier

1. INTRODUCTION

Progress in communication and information technology is very rapid and plays an important role

in human life in fulfilling needs that must be met. Technological advances that utilize the internet have an impact on human life patterns, such as how to interact with others, exchange information, and

million Instagram users in July 2023. The emergence of *cyberbullying* behavior is because Instagram provides comment facilities for its users on someone's video content so that other users can easily make hate speech or hate comments [7].

To reduce the occurrence of *cyberbullying* cases, bullying comments on Instagram must be classified to determine whether they fall into the bullying or non-bullying category. *Naïve Bayes Classifier* (NBC) and *K-Nearest Neighbor* (KNN) are two popular classification methods[8]. This research uses the *Naïve Bayes Classifier* (NBC) and *K-Nearest Neighbor* (KNN) algorithms which will compare the accuracy of the two methods after going through the preprocessing stage. The application of the *Naïve Bayes* method is relatively easy so it is widely used in research. The *Naïve Bayes Classifier* method can also be applied to different domains. The *K-Nearest Neighbor* (KNN) method is a margin increasing method that falls into two different classes. According to previous research, the *K-Nearest Neighbor* (KNN) and *Naïve Bayes* methods are the classifications with the highest accuracy [9].

Research using these two methods has been carried out by several researchers. According to research by Shaikh Zainuddin Nahdlatul Wathan Anjani regarding opinions about the public's response to PPPK teachers, using the *Naïve Bayes Classifier* and *K-Nearest Neighbor* algorithms. The results of sentiment analysis on 519 data with an accuracy result of 75.53% for the *Naïve Bayes Classifier* while the *Naïve Bayes Classifier* algorithm with an *accuracy* result of 73.41%[10]. Midhat Pasha Firdaus' research in sentiment analysis of the 2024 election predictions uses the *Naïve Bayes Classifier* and *K-Nearest Neighbor* algorithms as methods. Using 4453 data with an accuracy value of 82.86% using NBC and 78.56% using KNN[11]. Then in Tamrizal's research regarding sentiment analysis towards BPJS Health on Twitter media using NBC and KNN with accuracy results of 80% for the NBC algorithm and 67% for the KNN algorithm[12]. This research is related research regarding *cyberbullying* Sentiment Analysis using the *Naïve Bayes Classifier* and *K-Nearest Neighbor* Algorithms with word weighting using TF-IDF.

2. RESEARCH METHODS

Explains the stages used by the author in conducting research. The research method flow can be seen in Figure 1.

Research for information that is currently being widely discussed by the public. With the limitations of conventional media such as radio, newspapers, TV. To express an opinion. Not closing social media is an appropriate substitute or alternative for information expression [1]. The use of social media is growing rapidly in Indonesia. According to Reportal 167 million in 2023. There are 153 million users over 18 years old, or 79.5% of the total population. Additionally, it is estimated that 78.5% of Internet users use at least one social media account. This value is predicted to continue to increase in the coming years. The existence of this technology makes it easy to interact and disseminate information to the general public. However, the benefits obtained from using information technology are often used by people for negative things, such as posting insulting words, hate speech, and spreading hoax news that can harm the party concerned [2].

The high number of internet users presents major threats and risks, such as *cyberbullying*. *cyberbullying* is the behavior of uploading or sending text or images that are rude and harmful using digital media or the internet [3]. According to the secretary general of APJII, Henri Kasyfi, almost half of the population of internet users in Indonesia have been victims of *cyberbullying*. This figure was obtained from a survey conducted among internet users from March to April 14 2023. The *results* obtained from the survey which had 5,900 samples stated that 49 percent admitted that they had experienced bullying on social media [4]. According to the secretary general of APJII, Henri Kasyfi, almost half of the population of internet users in Indonesia have been victims of *cyberbullying*. This figure was obtained from a survey conducted among internet users from March to April 14 2023. The *results* obtained from the survey which had 5,900 samples stated that 49 percent admitted that they had experienced bullying on social media. The basic understanding of bullying involves the intimidator and the victim being bullied both physically and verbally, which can be done directly or indirectly[5].

Cyberbullying is often carried out on social media platforms. Instagram is a social media application that is accessed by almost all groups. Through Instagram, someone can upload photos or videos, publish them, the friendship system on Instagram uses the terms follower (*follower*) and following (*people who are followed*) [6]. There are like and comment features so that people can give appreciation in the form of likes or comments on uploaded photos. According to the We Are Social report, the number of global Instagram users will reach 6 billion in July 2023. The number of Instagram users in Indonesia is the 4th largest in the world, with around 103.3 million users. Meanwhile, India still occupies the top position with 332.15

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.

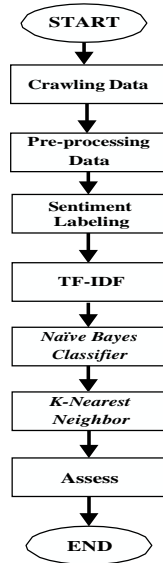


Figure 1. Research Method Flow

The research method is a flow for carrying out the stages in the research that will be carried out as shown in Figure 1.

2.1 Crawling Data

Crawling or data collection is collecting data for datasets [13]. *Crawling* is a technique used to collect information on the web. *Crawling* works automatically, where the information collected is based on keywords provided by the user. The tool used to crawl is called a crawler. Crawler is a program that is programmed with a certain algorithm, so that it can scan web pages, according to the web address or keywords provided by the user. When scanning, the crawler will read the existing text, hyperlinks and various tags used on the web page. Based on this information, the crawler will index the information or save the information in a file or database.

2.2 Pre-processing Data

Text Preprocessing is a process for improving the quality of text or selecting text data to eliminate noise [14]. *Pre-processing Data* is a data mining technique that involves transforming raw data into an easy-to-understand format. The data pre-processing step is necessary to resolve several types of problems including noisy data, data redundancy, missing data values, etc.

The data pre-processing steps are tokenization, case folding, stemming, filtering, and labeling.

1. Tokenization

Tokenization is the process of separating a series of characters based on space characters, and perhaps at the same time the process of removing certain characters, such as punctuation marks, is also carried out [15].

2. Case Folding

Case folding is the process of changing all the letters in a document or sentence to lowercase. The folding case is used to make searching easier. Not all data is consistent in the use of capital letters [16].

3. Stemming Words

That have been changed to lowercase need to be checked. Stemming is used to standardize words so as to reduce the list of words in the training data [17].

4. Filtering

Filtering/eliminating stopwords has many advantages, namely reducing space in the term index table by up to 40% or more. The stopwords removal process is the process of removing terms that have no meaning or are irrelevant [18].

5. Labelling

Labeling comes from the word label. Label means a character or set of characters used to identify a variable or part of data or a file. There are 2 (two) labeling processes, namely labeling the token with the word exaggeration and the word negation [19].

Word weighting is the process of giving weight to each word contained in a document. In searching for ranking information based on word frequency, one of the most popular methods is the TFIDF (Term Frequency - Inversed Document Frequency) method.

In the TF-IDF method, Term Frequency focuses more on terms that frequently appear in a document, while Inverse Document Frequency focuses more on giving low weight to terms that appear in many documents.

Naive Bayes classification is a classification that is supervised learning because it has a supervisor (a human who manually classifies the data used in training) as a teacher in the learning process. Apart from that, Naive Bayes performance has a short classification time, thereby speeding up the sentiment analysis system process. In this research, testing was used using 2 classes (negative and positive) [20].

2.3 Text Mining

Text mining is a mining process carried out by a computer to obtain something new, previously unknown, or to rediscover implicitly implied information. The findings come from information extracted automatically from different text data sources. Text Preprocessing is part of Text Mining which is done to remove noise in sentences. Text Preprocessing aims to avoid imperfect data, interference with data and non-consumable data [21].

2.4 Sentiment Labeling

Sentiment labeling is giving a sentiment value to a text which can be positive, negative or neutral. This labeling can be done using various methods, including manually, namely inputting the sentiment

value contained in the text yourself, then using the TF-IDF function, and with Python it can also be done using the TextBlob library.

2.5 TF-IDF Weighting

TF-IDF is a method that is an integration between term frequency (TF) and inverse document frequency (IDF). Term Frequency is calculated using Equation with the i th term frequency being the frequency of appearance of the i th term in the j th document. Inverse Document Frequency (IDF) is the logarithm of the ratio of the number of all documents in the corpus to the number of documents that have the term in question as written mathematically in Equation . The value is obtained by multiplying the two which is formulated in Equation [22].

$$idf_t = \log \frac{N}{df_t} \quad (1)$$

$$w_{t,d} = tf_{t,d} \times idf_t \quad (2)$$

2.6 Naïve Bayes Classifier (NBC)

Naïve Bayes is a classification method using simple probability which is rooted in Bayes' Theorem and has a high assumption of independence from each condition or event [23]. This method is part of the Bayes method used in text classification, based on a simplification model that attribute values are conditionally independent if the output value is given. The Naïve Bayes method is widely used in classification techniques on Twitter. This method is also used in text mining in sentiment analysis by predicting probabilities based on previous data. Naïve Bayes cannot detect images, but can only detect text and numbers. To calculate probability in this method, the Bayes theorem approach is used [24].

2.7 K-Nearest Neighbor (K-NN)

K-Nearest Neighbor (K-NN) is a simple method that is easy to implement, the data used has labels making it easier to group the process into the most appropriate classes and has the advantage of being able to classify data using training data and test data and having the ease of translating results and accuracy of predictions by accurately selecting the closest k value first [25].

3. RESULTS AND DISCUSSION

3.1 Crawling Data

The data that has been obtained using the data crawling technique is 400 comment data. The results of data crawling can be seen in Figure 2 .

| Id | Instagram Comment Text |
|-----|--|
| 0 | 1 <USERNAME> TOLOL!! Gak ada hubungan nya kegug... |
| 1 | 2 Geblek lo tata...cowo bgt dibela2in balikan..... |
| 2 | 3 Kmrrn termewek2 skr lengket lg duhhhh kok labil ... |
| 3 | 4 Intinya kalau kesel dengan ATT nya, gausah ke ... |
| 4 | 5 hadewwwww perempuan itu lg!!!!sakit jiwa,knp ha ... |
| ... | ... |
| 395 | 396 Bangga sama suami yg selalu ingat istri disela... |
| 396 | 397 Apaoun pekerjaannya yg penting halai u tuk men... |
| 397 | 398 Gojek itu mayoritas pengangguran yang lama gak ... |
| 398 | 399 <USERNAME> aslinya cantik dan ayu loh mbak kr... |
| 399 | 400 <USERNAME> suami saya seumuran sama saya mba... |

Figure 2. Data Crawling Results

The image above is the result of comment data obtained from Instagram social media.

3.2 Pre-processing Data

This stage is carried out after we have carried out the data crawling stage. Pre-processing results can be seen in table 1.

Table 1. Results Pre-processing Data

| Alur Text Preprocessing | Before | After |
|-------------------------|---|--|
| Cleaning | @fahrianalimbo ng tolol!! gak ada hubungan nya keguguran dgn pake hijab syar'i yg lo bilang bayi nya kepanasan didalem gak ada hubungan nya woyyyy!! otak sama jempol lo gak singkron sih ya jadinya asal nulis komentar! | tolol gak ada hubungan nya keguguran dgn pake hijab syar'i yang lo bilang bayi nya kepanasan didalem gak ada hubungan nya woyyyy otak sama jempol lo gak singkron sih ya jadinya asal nulis komentar |
| Tokenization | tolol hubung gugur pakai hijab syar i bilang bayi panas dalam hubung otak jempol singkron tulis komentar | 'tolol', 'hubung', 'gugur', 'pakai', 'hijab', 'syar', 'i', 'bilang', 'panas', 'dalam', 'hubung', 'otak', 'jempol', 'singkron', 'tulis', 'komentar' |
| Case Folding | @fahrianalimbo ng TOLOL!! Gak ada hubungan nya keguguran dgn | @fahrianalim bong tolol!! gak ada hubungan nya keguguran dgn |

Hak Cipta Dilindungi Undang-Undang

Stemming

Stopword Removal

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.

pake hijab syar'i yang lo bilang bayi nya kepanasan didalam gak ada hubungan nya woyyyy!! Otak sama jempol lo gak sinkron sih ya jadinya asal nulis komentar!

tolol hubungan keguguran pakai hijab syar'i bilang bayi kepanasan didalam hubungan otak jempol sinkron tulis komentar

tolol tidak ada hubungan nya keguguran dengan pakai hijab syar'i yang kamu bilang bayi nyakepanasan didalam tidak ada hubungan nya woi otak sama jempol kamu tidak sinkron sih ya jadinya asal tulis komentar

pake hijab syar'i yang lo bilang bayi nya kepanasan didalam gak ada hubungan nya woyyyy!! otak sama jempol lo gak sinkron sih ya jadinya asal nulis komentar!

tolol hubung gugur pakai hijab syar i bilang bayi panas dalam hubung otak jempol sinkron tulis komentar

tolol hubungan keguguran pakai hijab syar'i bilang bayi kepanasan didalam hubungan otak jempol sinkron tulis komentar

From the table above, you can see the Text Preprocessing results in table 1. Datasets that have passed the data preprocessing stage are uniform, structured and the symbols in the text have disappeared. This is because when you enter the next stage, namely the classification stage, it is more optimal in calculations.

3.3 Sentiment Labeling

Sentiment labeling uses the Python programming language by creating a negative and positive dictionary in Indonesian and creating a system. In this labeling we can see the differences between words that include positive and negative things. The results of sentiment labeling can be seen in Figure 3 below.

| Id | Sentiment | Instagram Comment Text |
|-----|-----------|--|
| 0 | 1 | negative <USERNAME> TOLOL!! Gak ada hubungan nya kegug... |
| 1 | 2 | negative Geblek lo lata...cowo bgt dibela2in balikan..... |
| 2 | 3 | negative Kmrn termewek2 skr lengket lg duhhh kok labil ... |
| 3 | 4 | negative Intinya kalau kesel dengan ATT nya, gausah ke ... |
| 4 | 5 | negative hadevwww permuan itu lg!!!!sakit jiwa.knp ha... |
| ... | ... | ... |
| 395 | 396 | positive Bangga sama suami yg selalu ingat istri disela... |
| 396 | 397 | positive Apaoun pekerjaannya yg penting halal u tuk men... |
| 397 | 398 | positive Gojek itu mayoritas pegangguran yang lama gak ... |
| 398 | 399 | positive <USERNAME> asinya cantik dan ayu loh mbak kr... |
| 399 | 400 | positive <USERNAME> suami saya seumuran sama saya mba... |

Figure 3. Sentiment Labeling Results

The labeling process aims to divide the class into 2 class parts, namely negative and positive classes. Where each sentence will contain the value of each word that contains the sentiment class.

3.4 TF-IDF

To complete the word weighting on the training data and test data, the next process is to calculate the TF-IDF value. The equation formula to get the word weighting results for each word (term) in each document can be known as an equation. Here's the process python TF-IDF weighting as with language programming is assisted by the Scikit learn library, TfidfVectorizer in figure 4.

```

Convert a collection of raw documents to a matrix of TF-IDF features
https://scikit-learn.org/stable/modules/generated/sklearn.feature_ext
...
from sklearn.feature_extraction.text import TfidfVectorizer

tf_idf = TfidfVectorizer(ngram_range=(1,1))
tf_idf.fit(X)
    
```

Figure 4. Script TF-IDF

Seen in Figure 4. The process of working on TfidfVectorizer is the process of transforming text documents into vectors. The results of the TF-IDF weighting can be seen in table 2 below.

Table 2. Results TF-IDF

| Term | Train Data | | | | | | Test Data | | |
|--------|------------|----|----|-------|----|----|-----------|-------|----|
| | D1 | D2 | D3 | D4 | D5 | D6 | U1 | U2 | U3 |
| tolol | 0 | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 |
| hubung | 0 | 0 | 0 | 0 | 0 | 0 | 1.24 | 0 | 0 |
| gugur | 0 | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 |
| pakai | 0 | 0 | 0 | 0.477 | 0 | 0 | 0.477 | 0.477 | 0 |
| hijab | 0 | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 |
| syar | 0 | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 |
| bilang | 0 | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 |
| bayi | 0 | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 |
| panas | 0 | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 |

1. Hak Cipta Dilindungi Undang-undang
 2. 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.

| | | | | | | | | | |
|----------|-------|---|---|---|---|-------|-------|-------|-------|
| dalam | 0 | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 |
| ok | 0 | 0 | 0 | 0 | 0 | 0 | 0.653 | 0.653 | 0 |
| jelepol | 0 | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 |
| sikron | 0 | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 |
| tulis | 0 | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 |
| komentar | 0 | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 |
| can | 1.412 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ong | 0.653 | 0 | 0 | 0 | 0 | 0 | 0 | 0.849 | 0 |
| nu | 0.849 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.849 |
| kualiti | 0.954 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mabuk | 0.954 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| teman | 1.526 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| libat | 0.954 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| bibir | 1.24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| coba | 0.954 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| bayangin | 0.954 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| sangup | 0.954 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| dasar | 0.954 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| laku | 0.954 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| joget | 0 | 0 | 0 | 0 | 0 | 0.954 | 0 | 0 | 0 |

3.3 Naïve Bayes Classifier

In the *Naïve Bayes Classifier* method in this research, as previously stated in the model stage, this research divides the training data and test data into 3 parts and is then processed using the *Naïve Bayes Classifier* method. The following are the results of dividing the training data dataset: test data can be seen in table 3.

Table 3. *Naïve Bayes* Results

| Data Latih : Data Uji | Accuracy | Presisi | Recall | f1-score |
|--------------------------|----------|---------|--------|----------|
| 70:30 | 91% | 91.5% | 91% | 91% |
| 80:20 | 95% | 95% | 95% | 95% |
| 90:10 | 93% | 92.5% | 92.5% | 92.5% |

The results of the confusion matrix using the *Naïve Bayes* method at a ratio of 90%:10% showed that there were 73 correct predictions for positive sentiment (*true positive*) and 7 data for correct predictions for negative sentiment (*true negative*).

3.4 K-Nearest Neighbor (K-NN)

The results of evaluating the dataset with the KNN algorithm with a K value by conducting experiments on the value k=3, with the partitioning validation evaluation model showing the best accuracy value with a k=3 value of 67% was obtained, as shown by the confusion matrix which is formed as follows data can be seen in table 3.

Table 4. *K-Nearest Neighbor* Results

| Data Latih : Data Uji | Accuracy | Presisi | Recall | f1-score |
|--------------------------|----------|---------|--------|----------|
|--------------------------|----------|---------|--------|----------|

| | | | | |
|-------|-------|-------|-------|-------|
| 70:30 | 87.5% | 29.2% | 33.3% | 31.1% |
| 80:20 | 88% | 28.9% | 33% | 31% |
| 90:10 | 90% | 30% | 30% | 31.5% |

The evaluation results using the KNN algorithm with a value of k=3 show an *accuracy* value of 67%, the confusion matrix formed from the application review data *processing* process.

3.5 The Final Result

This stage is the stage to determine which method results with the best accuracy. For a comparison of the two methods, see the picture.

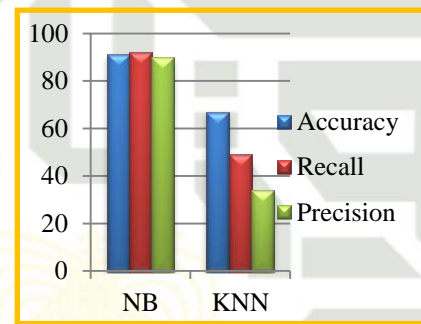


Figure 5. Comparison of Two Methods

In this research, the results of calculations using the first three methods using the Partitioning model, the results using the *Naïve Bayes* method obtained an accuracy value of 91.25%, a recall value of 92% and a precision value of 90% and calculations using the KNN method had an accuracy value of 67%, a recall value of 49% and a precision value of 34%.

4. DISCUSSION

Of the three related studies that have been explained in the introduction, researchers will carry out the same research, but with different research cases and testing with different data on the selected algorithm, with the aim of comparing which two algorithms are more accurate in determining the accuracy of *cyberbullying* levels. .

The researchers conducted sentiment analysis research to find out a person's opinion or response to *cyberbullying*, by carrying out an algorithm classification to find out the accuracy value regarding the selected topic data, to find out the opinion or response of the Indonesian people, especially on the topic of *cyberbullying* from Instagram social media comments, the researchers conducted analysis research sentiment using the *Naïve Bayes Classifier* and *K-Nearest Neighbor* algorithms.

In this case, sentiment analysis is very helpful to find out the responses or opinions of the Indonesian people regarding how to make positive and negative comments so that bullying or *cyberbullying* does not occur. So, in determining

Whether comments are positive or negative, the author takes data from Instagram social media comments. In determining the results, the author uses two algorithms to compare which algorithm is more accurate in identifying *Cyberbullying*. The results can be proven by the percentage of sentiment classes in Figures 5 and 6. From these two algorithms, there is an algorithm that is very effective in finding out how big the percentage accuracy of *cyberbullying* comments is.

CONCLUSION

Based on the results of the tests that have been carried out, several conclusions can be drawn in analyzing *cyberbullying* sentiment in Instagram comments using the Naïve Bayes Classifier and K-Nearest Neighbor methods. What is concluded from this research is as follows :

1. TF-IDF helps the *Naïve Bayes Classifier* and *K-Nearest Neighbor* algorithm in weighting words in text analysis.
2. The *Naïve Bayes Classifier* and *K-Nearest Neighbor* algorithms can be used to analyze *cyberbullying* sentiment in the Instagram comments column. The classification results are in the form of positive classes and negative classes which are divided into positive sentiment and negative sentiment.
3. Testing was carried out with two algorithms with the same data, the results of the first method were *Naïve Bayes Classifier* with an *accuracy* value of 91.25% with an average value for each sentiment class of 90% *precision* and 92% *recall* and calculations using the KNN method with an *accuracy* value of 67 %, *recall* value 49% and *precision* value 34%.
4. From these two methods, the author found the method with the best accuracy in *cyberbullying* sentiment analysis, namely the *Naïve Bayes Classifier* method with an *accuracy* value of 91.25% with an average value for each sentiment class of 90% *precision* and 92% *recall*.
5. *Text Preprocessing* has a big influence on classification to increase *accuracy* in sentiment analysis.

DAFTAR PUSTAKA

- [1] W. purbo Onno, *Text mining analisis medsos kekuatan brand & intelegen di internet*, 1st ed. yogyakarta: andi (anggota IKAPI), 2019
- [2] Rusdianan, D., dan Rosiyadi, D. *Analisa Sentimen Terhadap Tokoh Publik Menggunakan Metode Naïve Bayes Classifier Dan Support Vector Machine*.

CESS (Journal of Computer Engineering System and Science), 4(2), 230– 235.2019

- [3] D. Darwis, E. S. Pratiwi, and A. F. O. Pasaribu, "Penerapan Algoritma Svm Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia," *J. Ilm. Edutic Pendidik. dan Inform.*, vol. 7, no. 1, pp. 1–11, 2020.
- [4] Sukesti, Restu. "Bianglala mimpi: antologi esai bengkel bahasa dan sastra Indonesia siswa SLTA Kabupaten Sleman." *J. Ilm. Pendidik Bahasa dan Sastra Indonesia.*, vol. 5, no. 2, pp 1-8, 2019.
- [5] Arofah, Zahrotul, Muhammad Roisul Basyar, and Anggraeny Puspaningtyas. "Strategi Penanggulangan Bullying pada Sekolah Menengah Pertama dalam Perspektif Collaborative Governance: Studi pada SMP Islam Tikung." *Mutiara: Jurnal Penelitian dan Karya Ilmiah.*, vol. 1, no 6, pp 227-235, 2023.
- [6] Saingo, Yakobus Adi. "Pemanfaatan Media Sosial Sebagai Sarana Pembentukan Karakter Kristiani Tingkat Sekolah Menengah Pertama." *Jurnal Shanan.*, vol. 6, no.1, pp 89-110, 2022.
- [7] Zulkarnain, Iskandar, et al. "Penerapan Digital Marketing Sebagai Strategi Komunikasi Pemasaran Usaha Kecil Dan Menengah Warga Depok Jaya." *DEDIKASI.*, vol. 1, no. 1, pp. 78–86, 2020.
- [8] Anam, M. Khairul, et al. "Penerapan Naïve Bayes Classifier, K-Nearest Neighbor (KNN) dan Decision Tree untuk Menganalisis Sentimen pada Interaksi Netizen dan Pemeritah." *MATRIK: J. Manaje, T. Informatika dan Rekayasa Komputer.*, vol. 21, no. 1, pp. 139-150, 2021.
- [9] Ilmawan, Lutfi Budi, and Muhammad Aliyazid Mude. "Perbandingan metode klasifikasi Support Vector Machine dan Naïve Bayes untuk analisis sentimen pada ulasan tekstual di Google Play Store." *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 154-161, 2020.
- [10] Hidayat, Fajar Nurrachmat, and Sugiyono Sugiyono. "Analisis Sentimen Masyarakat Terhadap Perekrutan Pppk Pada Twitter Dengan Metode Naive Bayes Dan Support Vector Machine." *Jurnal Sains dan Teknologi.*, vol. 5, no. 2, pp. 665-672, 2023.
- [11] Zain, Haekal Hilmi, Rolly Maulana Awangga, and Woro Isti Rahayu. "Perbandingan Model Svm, Knn Dan Naïve Bayes Untuk Analisis Sentiment Pada Data Twitter: Studi Kasus Calon Presiden 2024." *JIM: Jurnal Ilmiah Mahasiswa Pendidikan Sejarah.*, vol.8, no. 3,

Hak Cipta © Ditujukan Kepada: UIN Suska 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

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.

[2] Puspita, Rani, and Agus Widodo. "Perbandingan Metode KNN, Decision Tree, dan Naïve Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS." *Jurnal Informatika Universitas Pamulang.*, vol.5, no. 4, pp. 646-654, 2021.

[3] A. R. Isnain et al., "Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm," *Universitas. Teknokr. Indonesia.*, vol. 2, no. 1, pp. 31– 37, 2021.

[14] A. P. Natasuwarna, "Seleksi Fitur Support Vector Machine pada Analisis Sentimen Keberlanjutan Pembelajaran Daring," *Techno.Com*, vol. 19, no. 4, pp. 437–448, 2020, doi: 10.33633/tc.v19i4.4044

[15] Singgalen, Yerik Afrianto. "Analisis Sentimen Konsumen terhadap Food, Services, and Value di Restoran dan Rumah Makan Populer Kota Makassar Berdasarkan Rekomendasi Tripadvisor Menggunakan Metode CRISP-DM dan SERVQUAL." *Building of Informatics, Technology and Science (BITS).*, vol. 4, no. 4. Pp. 1899-1914., 2023.

[16] Alita, D., & Isnain, A. R. Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier. *Jurnal Komputasi.*, vol. 8, no. 2, pp. 50-58, 2020.

[17] Hickman, L., Thapa, S., Tay, L., Cao, M., & Srinivasan, P. Text preprocessing for text mining in organizational research: Review and recommendations. *Organizational Research Methods.*, vol. 25, no. 1, pp. 114-146, 2022.

[18] Rianto, et al. "Improving the accuracy of text classification using stemming method, a case of non-formal Indonesian conversation." *Journal of Big Data .*, vol. 8, pp. 1-16, 2021.

[19] Siregar, Amril Mutoi, and Maulana Abdur Rofik. "Mengukur Tingkat Kepuasan Mahasiswa Dalam Pembelajaran Dengan Naïve Bayes." *Prosiding Konferensi Nasional Penelitian Dan Pengabdian Universitas Buana Perjuangan Karawang.*, vol. 1, no. 1, pp. 1-19. 2021.

[20] Juniarsih, Sri, Eva Faja Ripanti, and Enda Esyudha Pratama. "Implementasi Naive Bayes Classifier pada Opinion Mining Berdasarkan Tweets Masyarakat Terkait Kinerja Presiden dalam Aspek Ekonomi." *JUSTIN (Jurnal Sistem dan Teknologi Informasi).*, vol. 8, no. 3, pp. 239-249, 2022.

[21] Sari, Fransiska Vina, and Arief Wibowo. "Analisis sentimen pelanggan toko online Jd. Id menggunakan metode Naïve Bayes Classifier berbasis konversi ikon emosi." *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer* 10.2 (2019): 681-686.

[22] Sumayah, Siti, Falentino Sembiring, and Wisuda Jatmiko. "Analysis of sentiment of Indonesian community on metaverse using support vector machine algorithm." *Jurnal Teknik Informatika (JUTIF).*, vol. 4,no, 1, pp. 143-150, 2023

[23] Watratan, Alvina Felicia, and Dikwan Moeis. "Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia." *Journal of Applied Computer Science and Technology.*, vol. 1, no. 1, pp. 7-14, 2020

[24] Prabowo, W. A., dan Wiguna, C. Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM. *Jurnal Media Informatika Budidarma.*, vol. 5, no. 1, pp. 149-156, 2021.

[25] Singgalen, Yerik Afrianto. "Analisis Sentimen Konsumen terhadap Food, Services, and Value di Restoran dan Rumah Makan Populer Kota Makassar Berdasarkan Rekomendasi Tripadvisor Menggunakan Metode CRISP-DM dan SERVQUAL." *Building of Informatics, Technology and Science (BITS).*, vol. 4, no. 4. Pp. 1899-1914., 2023.



LAMPIRAN A
EMAIL ACCEPTED

© Hak cipta UIN Suska Riau



Jutif | New notification from Jurnal Teknik Informatika (Jutif)

Pesan

Silang Ratmana <jutif.ft@unsoed.ac.id>

Rab, 8 Mei 2024 pukul 10.30

Balas ke: JUTIF UNSOED <jutif.ft@unsoed.ac.id>

Kepada: Fitri Anisa Nirmala <12050322148@students.uin-suska.ac.id>You

You have a new notification from Jurnal Teknik Informatika (Jutif):

You have added to a discussion titled "Letter of Acceptance (LoA)" regarding the submission "A CYBERBULLYING SENTIMENT ANALYSIS OF INSTAGRAM COMMENTS USING NAÏVE BAYES AND K-NEAREST NEIGHBOR ALGORITHM METHODS".

Link: <https://www.jutif.if.unsoed.ac.id/index.php/jurnal/authorDashboard/submission/1997> JUTIF

UNSOED

[Journal of Informatics Engineering \(Jutif\)](#)

State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Diindungi Undang-Undang

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

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

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

LAMPIRAN B

EMAIL REVISION



[jutif] Editor Decision

1 pesan

Lasmedi Afuan <jutif.ft@unsoed.ac.id>

Jum, 3 Mei 2024 pukul 13.46

Kepada: Fitri Anisa Nirmala <12050322148@students.uin-suska.ac.id>

Fitri Anisa Nirmala:

We have reached a decision regarding your submission to Jurnal Teknik Informatika (Jutif), "A CYBERBULLYING SENTIMENT ANALYSIS OF INSTAGRAM COMMENTS USING NAÏVE BAYES AND K-NEAREST NEIGHBOR ALGORITHM METHODS: Introduction, Research Methods, Results And Discussion, Discussion, Crawling Data, pre-processing Data, Naive Bayes Classifier (NBC), K-Nearest Neighbor (K-NN)".

Our decision is: Revisions Required

In addition, several things to consider from the editor's side are as follows :

1. Please **FOLLOW** JUTIF's template. You can download it at <http://jutif.if.unsoed.ac.id/index.php/jurnal/about/submissions> after that, you can copy your paper into template.
2. Please use table and figure format according to JUTIF's Template. The resolution of figures must be clear, can be seen and **more than 300 dpi**. You can make it bigger into 1 column if necessary.
3. Every figure and table **MUST BE** cited and must be explained in paragraph, ex : "Research method can be seen in Figure 1", "Table 1 showed XYZ."
4. Section "Discussion" is **MANDATORY**. If your article did not contain Discussion, please add it.
5. **MINIMUM** references in JUTIF is 15 primary references (Journal/Conference) from 5 years latest. Please add reference if necessary. In JUTIF, Reference using IEEE format. Please adjust in your reference.
6. Please Correct the manuscript **according to the review** from the reviewer or editor. Comments from reviewers that need to be corrected for the manuscript has been moved at the bottom. Reviewers also give a file that need to be corrected. Please login into Online Journal Systems (OJS) Jurnal Teknik Informatika and download that file.
7. Please give mark the revised results on the revised manuscript sent with **YELLOW HIGHLIGHTS** to the corrected sentence or sentences added of revision fulfillment from the Reviewer.
8. Revisions can be sent via the "File upload" in the "revision" section of the JUTIF's OJS with **maximum 3 weeks** after this email is sent. Please when upload the Revised File, do not overwrite the previous revised file.

Thank You.

Lasmedi Afuan
Informatics, Universitas Jenderal Soedirman
lasmedi.afuan@unsoed.ac.id

Reviewer A:

Recommendation: Revisions Required

Title

Judul baik, karena mengandung informasi tentang objek penelitian, metode dan tujuan penelitian.

Abstract

Abstrak cukup, karena mengandung permasalahan, tujuan, metode dan hasil penelitian.

Saran :

- Silahkan kata asing ditulis italic.

Introduction

Pendahuluan dan Latar Belakang masih kurang, karena menyajikan permasalahan atau pentingnya penelitian dan memuat penelusuran penelitian dan diperkuat rujukan dengan jumlah yang kurang (minimal 15 rujukan).

Saran :

- Silahkan ditambahkan penelusuran penelitian dan diperkuat rujukan dengan jumlah yang cukup (minimal 15 rujukan).
- Penyajian rujukan sebaiknya ditulis urut mulai 1 – 15.

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



Hak Cipta Diindungi Undang-Undang

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

Method

Metode kurang, karena menyajikan metode yang digunakan sesuai tahapan penelitian yang digunakan namun belum lengkap.
 Saran :
 - Silahkan bagian metode menyajikan alur penelitian dengan mengakomodir metode yang digunakan dan dilengkapi dengan flowchart alur penelitian.
 - Penyajian persamaan silahkan disesuaikan dengan format JUTIF

Results

Hasil cukup, karena menyajikan hasil penelitian, namun belum sesuai dengan tahapan penelitian yang dilakukan.
 Saran :
 - Silahkan gambar dan tabel yang disajikan dirujuk pada deskripsi yang disajikan.
 - Penyajian deskripsi pada bagian hasil sebaiknya tidak melampirkan rujukan karena merupakan hasil penelitian yang dihasilkan oleh peneliti sendiri, misalnya pada bagian TF-IDF masih menggunakan rujukan [20]

Discussions

Diskusi kurang, karena Bagian Diskusi belum menyajikan pendapat penulis dan membandingkan hasil penelitian yang dijalankan dengan penelitian lain yang serupa.
 Saran :
 - Silahkan bagian Diskusi menyajikan pendapat penulis dan membandingkan hasil penelitian yang dijalankan dengan penelitian lain yang serupa.

Conclusion

Kesimpulan kurang, karena belum menyimpulkan hal-hal yang disajikan dan dibahas di bagian hasil.
 Saran:
 - Silahkan ditambahkan kesimpulan menyimpulkan hal-hal yang disajikan dan dibahas di bagian hasil

References

Daftar Pustaka kurang, karena Jumlah sitasi minimal 15 dengan kurang 80% sitasi berasal dari Jurnal/seminar 5 tahun terakhir.
 Saran :
 - Silahkan ditambahkan sitasi minimal 15 dengan 80% sitasi berasal dari Jurnal/seminar 5 tahun terakhir.
 - Penulisan Daftar Pustaka harus dengan gaya IEEE

Overall Comment

Secara umum artikel baik, namun perlu melakukan revisi sesuai saran

Reviewer B:
 Recommendation: Revisions Required

Title

The title is in accordance with the provisions of writing and provides information on the object and method of research.

Abstract

The abstract is in accordance with the provisions of writing and provides information on problems, methods, and research summary result.

Introduction

The background already presents the problem, the purpose of the research, advice related to background, and the importance of research to be conducted.

Method

The method is in accordance with the object of research and is in accordance with the writing

Results

The results of the research have been presented in detail and arranged according to the format of writing





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.

Discussions

The discussion has been arranged according to the format of writing and personal getters have been included for this study

Conclusion

The conclusion has been described in detail and in accordance with the terms of writing.

References

There are several bibliography lists that do not include the year of publication

Overall Comment

Need improvement in the reference section.

Reviewer C:

Recommendation: Revisions Required

Title

Judul sudah lengkap dan merepresentasikan Artikel

Abstract

Abstrak sudah memuat latar belakang masalah, tahapan riset, mencantumkan hasil dan kesimpulan.

Introduction

Pendahuluan sudah lengkap beserta referensi 5 tahun terakhir terbaru.

Method

Metode penelitian sudah lengkap diberikan penjelasan untuk setiap tahapan.

Results

Hasil dan pembahasan sudah lengkap, hanya saja mohon gambar diperjelas

Discussions

Pada bagian hasil dan pembahasan sudah detail, hanya saja untuk tabel perlu disesuaikan. Ada beberapa gambar yang perlu diperbaiki, kurang bisa terbaca

Conclusion

Kesimpulan sudah menjawab dengan tujuan penelitian.

References

Daftar Pustaka sudah update mohon untuk dapat mensitasi artikel yang di JUTIF yang sesuai dengan Topik saudara.

Overall Comment

Secara keseluruhan sudah sesuai format jutif, beberapa harus di revisi sedikit

Jurnal Teknik Informatika (Jutif)



DAFTAR RIWAYAT HIDUP

Fitri Anisa Nirmala lahir di Desa Kabun, pada tanggal 16 Februari 2002. Peneliti merupakan anak dari Bapak Jumadi dan Ibu Nurrohmah. Peneliti adalah anak ketiga dari empat bersaudara. Pada tahun 2007 peneliti memulai pendidikan dengan masuk TK Ananda di Desa Kabun dan lulus pada tahun 2008. Lalu melanjutkan pendidikan Sekolah Dasar di SD Negeri 007 Kabun. Peneliti menyelesaikan pendidikan Sekolah Dasar pada tahun 2014. Setelah menyelesaikan pendidikan Sekolah Dasar peneliti melanjutkan pendidikan tingkat SLTP di SMP Negeri 001 Kabun yang selesai pada tahun 2017. Peneliti melanjutkan pelajaran ke tingkat SLTA di SMK LPMD Kabun. Setelah menyelesaikan pendidikan di SMK LPMD Kabun pada tahun 2020, peneliti pun melanjutkan pendidikan dengan menjadi mahasiswa Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Pada tahun 2022, peneliti telah menyelesaikan kerja praktek yang bertempat di Kamari Transportasi Online Karya Anak Nagari Pasaman Timur, kemudian selanjutnya pada tahun 2023 peneliti juga telah menyelesaikan Kuliah Kerja Nyata (KKN) di Desa Padang Luas, Kecamatan Langgam, Kabupaten Pelalawan. Dan peneliti telah menyelesaikan kuliah Strata satu (S1) dengan mengambil judul tugas akhir “Cyberbullying Sentiment Analysis Of Instagram Comments Using Naïve Bayes Classifier And K-Nearest Neighbor Algorithm Methods” pada tahun 2024.



© Hak

Suska Riau

State Islamic University of Sultan Syarif Kasim 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.