

Analisis Sentimen Pengguna Indihome dengan Metode Klasifikasi Support Vector Machine (SVM)

Muhammad Nur Akbar^{1*}, Nur Annisa Safitri Yusuf², Nasrullah³, Mubarak⁴

^{1,2,3,4}Jurusan Teknik Informatika, Universitas Islam Negeri Alauddin Makassar, Indonesia

¹muhnurakbar@uin-alauddin.ac.id, ²60200118005@uin-alauddin.ac.id, ³60200118033@uin-alauddin.ac.id,

⁴60200118087@uin-alauddin.ac.id

Informasi Artikel

Article history:

Diterima Des 29, 2021

Revisi 1 Des 31, 2021

Revisi 2 Jan 2, 2021

Dipublis Jan 30, 2022

Kata Kunci:

Setimen Analisis,
Twitter,
SVM

ABSTRACT

Indonesia Digital Home (IndiHome) is a communication service provider from PT Telekomunikasi Indonesia (Telkom) that provides several communication and data service packages such as internet, home telephone and cable television (Use TV & IP TV) which implements copper and fiber optic cable services. Currently, IndiHome is implementing a 100% fiber service replacement for all customers in order to produce high data speeds and more reliable services. However, the fact is that fiber optic services often receive complaints from customers due to weather and other factors. It was recorded that in 2020 internet users in Indonesia reached 196.7 million people or 73.7% million of the population and around 51.2% were social media users (Kompas.com, 2020). One of the social media with 6.43 million active users is Twitter. Twitter then became a medium for channeling opinions regarding a service, including the services provided by Indihome. Based on this, a method is needed, namely sentiment analysis to understand whether the opinion is negative or positive. The Support Vector Machine (SVM) is used to create a classification model for sentiment analysis of IndiHome service users' opinions on Twitter with an accuracy of 91.3%.

*Koresponden Author:

Muhamad Nur Akbar,
Jurusan Teknik Informatika,
Universitas Islam Negeri Alauddin Makassar,
Jl. H.M. Yasin Limpo No. 36 Samata, Kab Gowa, Sulawesi Selatan, Indonesia.
Email: muhnurakbar@uin-alauddin.ac.id



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

1. PENDAHULUAN

Seiring dengan kemajuan zaman serta kebutuhan manusia semakin meningkat terutama di bidang teknologi, informasi dan komunikasi, sehingga mendorong konsumen untuk meningkatkan intensitas penggunaan teknologi dalam kebutuhan sehari-hari. Internet merupakan salah satu perkembangan teknologi informasi dan komunikasi yang menawarkan sisi kecanggihan, kepraktisan, dan kemudahan. Pada tahun 2020 pengguna internet di Indonesia mencapai 196,7 juta orang atau 73,7% juta dari populasi dan sekitar 51,2% merupakan pengguna media sosial [1]. Salah satu media sosial dengan pengguna aktif sebesar 6,43 juta pengguna adalah twitter. Twitter merupakan media sosial yang digunakan untuk mencari informasi tentang bisnis, hiburan, ekonomi, politik, dan lainnya. Dengan banyaknya jumlah tweet yang dipublikasikan melalui twitter, tweet tersebut dapat

mengandung opini pengguna mengenai sebuah objek, objek tersebut dapat berupa kejadian disekitar masyarakat seperti suatu produk atau layanan.

Pandangan masyarakat atau umpan balik selalu terbukti menjadi sumber daya paling penting dan berharga bagi perusahaan. Sehingga berdasarkan tweet yang mereka publikasi, dapat dilihat gambaran mengenai opini pengguna melalui proses penggalian informasi lebih lanjut dari tiap tweet yang dipublikasi, akan tetapi jumlah data tersebut banyak sehingga dibutuhkan sebuah metode untuk dapat mengubah data tersebut menjadi sebuah informasi atau sebuah pengetahuan.

Suatu metode yang dapat digunakan untuk mewujudkannya yaitu dengan menggunakan analisis sentimen. Analisis sentimen ialah suatu metode dalam memahami, mengekstraksi data sentimen yang biasanya akan dikategorikan berdasarkan polaritasnya, apakah positif atau negative. Saat ini analisis sentimen merupakan salah satu cara untuk mengetahui atau mendapatkan informasi mengenai penilaian suatu brand, rating, dan penyaringan opini untuk membantuk perusahaan ataupun masyarakat. Proses analisis sentimen mencakup teks ulasan, forum, tweet atau blog dengan preprocessing yang mencakup tokenizing, stopword, penghapusan, stemming, identifikasi sentimen dan klasifikasi sentimen [2] [3] [4]. Fokus analisis sentimen adalah pendapat yang menyatakan positif atau negatif sehingga analisis sentimen dapat juga dikatakan sebagai opinion mining [5]. Twitter adalah media komunikasi yang umumnya digunakan pengguna internet untuk menyatakan ekspresi, menyebarkan berita dan memberi kabar. Tweet merupakan salah satu layanan yang disediakan oleh twitter, pengguna menuliskan opini, pendapat, kabar ataupun memasukkan link atau gambar dengan menggunakan akun twitternya [6]. Dengan banyaknya jumlah pengguna twitter, menjadikan twitter sebagai sumber data teks yang dapat memberikan informasi yang lebih bermanfaat [7].

Ada banyak penelitian yang menggunakan twitter sebagai sumber data primer diantaranya adalah penelitian analisis sentimen mengenai Netflix yang diblokir Telkom dengan menggunakan metode *Support Vector Machine* oleh Rizqa. Dimana penelitian tersebut memiliki *Accuracy* sebesar 86.04%, *Precession* 86.49%, *Recall* 92.44%, *F Measure* sebesar 89.36% [8]. Penelitian analisis sentimen yang dilakukan oleh Pangestu tahun 2019 dengan objek opini tentang partai politik di Indonesia. Penelitian tersebut bertujuan untuk melakukan klasifikasi sikap politik terhadap partai politik Indonesia dengan metode *Support Vector Machine* (SVM) yang memiliki hasil *accuracy* sebesar 86% dan pengujian dengan *10-Fold Cross Validation* rata-rata *accuracy* sebesar 71% dengan *error rate* sebesar 29% [9].

Berdasarkan latar belakang diatas, serta berpedoman pada penelitian sebelumnya, maka didapat rumusan masalah, yaitu bagaimana mendapatkan model klasifikasi sentimen menggunakan *Support Vector Machine* (SVM). SVM ditemukan oleh Vladimir Vapnik, SVM adalah salah satu metode klasifikasi yang dilakukan dengan garis pembatas (hyperlane) yang memisahkan antara opini positif dan negatif. Suatu garis pembatas yang baik adalah yang memiliki jarak terbesar ke titik data pelatihan terdekat dari setiap kelas, karena pada umumnya semakin besar margin, semakin rendah error generalisasi dari pemilah. Margin adalah jarak dari suatu titik vektor di suatu kelas terhadap hyperplane [9]. Ruang lingkup dari penelitian ini yaitu 452 data tweet dari 1.000 data yang telah sortir secara manual, dimana data yang digunakan hanya yang menyebutkan/mention akun Indihome (@IndiHome), tweet yang digunakan berbahasa Indonesia dan metode klasifikasi yang diterapkan adalah *Support Vector Machine* (SVM). Adapun tujuan dari penelitian ini adalah untuk dapat pemodelan klasifikasi sentimen dengan *Support Vector Machine* (SVM), untuk mengetahui seberapa besar akurasi yang dihasilkan oleh metode SVM, serta untuk mengetahui seberapa puas pengguna layanan Indihome berdasarkan twitter.

2. METODE PENELITIAN/ALGORIMA

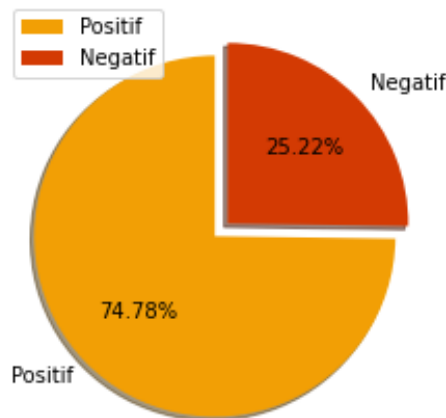
Data yang digunakan pada penelitian merupakan data *crawling* yang diambil dari twitter dengan menggunakan *tools orange*. Data yang diambil berdasarkan kumpulan tweet yang ditujukan kepada akun twitter IndiHome, yakni dengan menggunakan kata kunci yang menyebut username @IndiHome. Data yang berhasil diambil adalah sebanyak 1000 data tweet dan disortir secara manual sehingga menjadi 452 data tweet. Contoh data tweet yang dikumpulkan dapat dilihat pada tabel 2.1.

Tabel 1. Contoh data tweets yang didapatkan dari hasil *crawling*

No	Text Tweets
1.	Indihome ngajak geludd
2.	inginku push rank tapi indihome berkata lain
3.	@bertanyarl indihome, 400an/bln cm ga worth si mnrt aku karena suka error

Langkah selanjutnya adalah pelabelan data pada kelas tweets menggunakan kategori positif dan negatif. Tanggapan positif pengguna IndiHome sebanyak 338 tweets sedangkan sentimen negatif sebanyak 114 tweets, yang ditunjukkan pada gambar 1 dan gambar 2.

Diagram Pie Data Analisis Sentimen



Gambar 1. Diagram pie data analisis sentimen

No	Tweets	Label
1	Indihome ngajak geludd	negative
2	inginku push rank tapi indihome berkata lain	negative
3	@bertanyarl indihome, 400an/bln cm ga worth si mnrt aku karena suka error	negative
4	ini Spotify berantem lagi sama indihome kah?!?! hadeh daritadi error mulu	negative
5	@bertanyarl indihome kata bapak gak smpe 500 kira kira 350-450 mnrtku worth sih dr	positive

Gambar 2. Contoh hasil pelabelan manual

Selajutnya yang dilakukan adalah *pre-processing*, *pre-processing* merupakan proses untuk mempersiapkan data mentah sebelum dilakukan proses lain. Proses Preprocessing pada penelitian ini dilakukan menggunakan *Jupiter Notebook* dengan memanfaatkan beberapa *library*. Adapun tahap *pre-processing* terdiri dari tahapan sebagai berikut:

1. *Text Cleaning*. Pada tahap ini merupakan prose membersihkan tweets, seperti menghapus *mention*, URL, hastag, tanda baca, angka, simbol. Proses *text cleaning* ini menggunakan *library* re dan *numpy*.

No	Tweets	Label	clean_tweets	
0	1	Indihome ngajak geludd	negative	Indihome ngajak geludd
1	2	inginku push rank tapi indihome berkata lain	negative	inginku push rank tapi indihome berkata lain
2	3	@bertanyarl indihome, 400an/bln cm ga worth si...	negative	indihome worth mnrt karena suka error
3	4	ini Spotify berantem lagi sama indihome kah?!?...	negative	Spotify berantem lagi sama indihome hadeh dari...
4	5	@bertanyarl indihome kata bapak gak smpe 500 k...	positive	indihome kata bapak smpe kira kira mnrtku wort...

Gambar 3. Contoh data yang telah di *text cleaning*

2. *Tokenizing*. Pada tahap ini *text* tweets akan diproses dengan merubah kalimat pada *text* tweets menjadi potongan kata. Proses tokenizing ini menggunakan *library* nltk.

No	Tweets	Label	clean_tweets	Tokenizing	
0	1	Indihome ngajak geludd	negative	Indihome ngajak geludd	[Indihome, ngajak, geludd]
1	2	inginku push rank tapi indihome berkata lain	negative	inginku push rank tapi indihome berkata lain	[inginku, push, rank, tapi, indihome, berkata,...]
2	3	@bertanyarl indihome, 400an/bln cm ga worth si...	negative	indihome worth mnrt karena suka error	[indihome, worth, mnrt, karena, suka, error]
3	4	ini Spotify berantem lagi sama indihome kah?!?...	negative	Spotify berantem lagi sama indihome hadeh dari...	[Spotify, berantem, lagi, sama, indihome, hade...]
4	5	@bertanyarl indihome kata bapak gak smpe 500 k...	positive	indihome kata bapak smpe kira kira mnrtku wort...	[indihome, kata, bapak, smpe, kira, kira, mnrtku, wort...]

Gambar 4. Contoh hasil *tokenizing*

3. *Stopwords*. Dimana tahap merupakan mengseleksi tweets dengan cara menghilangkan kata-kata yang tidak penting atau tidak memiliki makna sesuai dengan kamus *stopword*. Pada proses *stopwords* ini kami menggunakan *library* nltk.corpus.

No	Tweets	Label	clean_tweets	Tokenizing	Stop Words	
0	1	Indihome ngajak geludd	negative	Indihome ngajak geludd	[Indihome, ngajak, geludd]	[Indihome, ngajak, geludd]
1	2	inginku push rank tapi indihome berkata lain	negative	inginku push rank tapi indihome berkata lain	[inginku, push, rank, tapi, indihome, berkata,...]	[inginku, push, rank, indihome]
2	3	@bertanyarl indihome, 400an/bln cm ga worth si...	negative	indihome worth mnrt karena suka error	[indihome, worth, mnrt, karena, suka, error]	[indihome, worth, mnrt, suka, error]
3	4	ini Spotify berantem lagi sama indihome kah?!?...	negative	Spotify berantem lagi sama indihome hadeh dari...	[Spotify, berantem, lagi, sama, indihome, hade...]	[Spotify, berantem, indihome, hadeh, daritadi,...]
4	5	@bertanyarl indihome kata bapak gak smpe 500 k...	positive	indihome kata bapak smpe kira kira mnrtku wort...	[indihome, kata, bapak, smpe, kira, kira, mnrt...]	[indihome, smpe, mnrtku, worth, drpda, orang, ...]

Gambar 5. Contoh hasil *stopwords*

4. *Stemming*. Pada tahap ini merupakan proses merubah semua tweet yang memiliki kata imbuhan menjadi kata dasar. Proses ini menggunakan *library* Nltk.stem.

```
0          [indihom, ngajak, geludd]
1          [inginku, push, rank, indihom]
2          [indihom, worth, mnrt, suka, error]
3  [spotifi, berantem, indihom, hadeh, daritadi, ...
4  [indihom, smpe, mnrtku, worth, drpda, orang, r...
Name: Stop Words, dtype: object
```

Gambar 6. Contoh hasil stemming

3. HASIL PENELITIAN DAN PEMBAHASAN

3.1. Text Mining

Text mining adalah salah satu penambangan informasi dengan cara mengeksplorasi dan menganalisis sejumlah besar data teks tidak terstruktur yang dapat mengidentifikasi konsep, pola, topik, kata kunci, dan atribut lainnya dalam data. *Text mining* telah menjadi lebih praktis bagi para ilmuwan data dan pengguna lain karena pengembangan platform data besar dan algoritma pembelajaran mendalam yang dapat menganalisis kumpulan data yang tidak terstruktur secara besar-besaran. Menganalisis teks membantu organisasi atau perusahaan menemukan *knowledge* atau potensi wawasan bisnis yang berharga dalam dokumen perusahaan, email pelanggan, log call center, komentar *survei verbatim*, posting jaringan sosial, catatan medis dan sumber data berbasis teks lainnya. Text mining merupakan suatu teknik yang digunakan untuk menggali informasi dan pengetahuan dari suatu teks [10]. [11] mendefinisikan text mining sebagai proses mengekstrak informasi dari teks yang digunakan untuk suatu tujuan tertentu. Di antara teknik yang digunakan dalam text mining yaitu mengukur asosiasi atau hubungan, network analysis, visualisasi dan analisis prediktif. Beberapa kelebihan text mining antara lain kecepatan analisis penggunaan teks, kata, dan interpretasi, serta kemudahan untuk memvisualisasikan hasil analisis sedangkan kekurangannya yaitu kuatnya subjektivitas dalam persepsi penilaian dalam mengukur sentimen suatu kata dalam teks [12].

Untuk dapat melakukan penambangan informasi atau *text mining* maka perlu dilakukan beberapa tahapan yang harus dilakukan untuk mengolah sumber data baik yang terstruktur, terstruktur sebagian dan yang tidak terstruktur dari beberapa sumber maka data-data tersebut perlu dilakukan proses awal atau di sebut sebagai *pre-processing text* yang bermaksud mengolah data awal yang masih bermacam-macam untuk dijadikan sebuah data teratur yang dapat dikenai atau diterapkan beberapa metode text mining yang ada [13].

3.2. TF-IDF

Term Frequency (TF) merupakan salah satu ukuran seberapa penting sebuah kata dalam sebuah dokumen. Namun ada kata-kata dalam dokumen yang muncul berkali-kali tetapi mungkin tidak penting. hal tersebut dapat diatasi dengan menambahkan kata tidak penting tersebut ke dalam daftar stopword (pada tahap *pre-processing*) dan menghapusnya sebelum melakukan analisis, akan tetapi ada kemungkinan bahwa beberapa dari kata-kata ini mungkin lebih penting dalam beberapa dokumen daripada yang lain.

Pendekatan lain adalah dengan melihat pada *Inverse Document Frequency* (IDF), yang mengurangi bobot untuk kata-kata yang umum digunakan dan meningkatkan bobot untuk kata-kata yang tidak terlalu banyak digunakan dalam kumpulan dokumen. Ini dapat dikombinasikan dengan frekuensi suku untuk menghitung tf-idf suku (dua besaran dikalikan bersama), frekuensi suku disesuaikan dengan seberapa jarang digunakan.

Statistik tf-idf dimaksudkan untuk mengukur seberapa penting sebuah kata bagi sebuah dokumen dalam kumpulan (atau korpus) dokumen, misalnya terhadap satu novel dalam kumpulan novel atau satu website dalam kumpulan website.

3.3. Algoritma Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan salah satu metode dalam *supervised learning* yang biasanya digunakan untuk klasifikasi (seperti *Support Vector Classification*) dan regresi (*Support Vector Regression*). Dalam pemodelan klasifikasi, SVM memiliki konsep yang lebih matang dan lebih jelas secara matematis dibandingkan dengan teknik-teknik klasifikasi lainnya. SVM juga dapat mengatasi masalah klasifikasi dan regresi dengan linear maupun non linear. Metode SVM memungkinkan perhitungan untuk masalah linear dengan menerapkan transformasi matematis untuk ruang belajar menggunakan fungsi kernel. Metode SVM memiliki konsep sentral dalam mengklasifikasikan data, yaitu menentukan *hyperplane* terbaik untuk memberi jarak atau pemisah antara dua kelas yang telah ditentukan.

Cara kerja SVM yaitu dengan dengan menemukan *hyperplane optimal* yang memberi jarak atau pemisah antar dua kelas. *hyperplane* yang mempunyai margin maksimum. Jarak antar titik data terdekat dengan *hyperplane* diklaim sebagai margin. Adapun *vector* pendukung sebagai titik terdekat dengan *hyperplane*. Adapun tahapan dalam metode *Support Vector Machine* adalah sebagai berikut:

1. Menentukan kata yang sering muncul dari tiap dokumen atau tweet yang digunakan.
2. Menentukan inisialisasi awal untuk nilai $\alpha=0.5$, $C=1$, $\lambda=0.5$, $\text{gamma}=0.5$ dan $\text{epsilon}=0.001$.
3. Menghitung matriks dengan rumus (2):

$$D_{ij} = y_i y_j (K(\vec{x}_i \cdot \vec{x}_j) + \lambda^2)$$

Dengan keterangan:

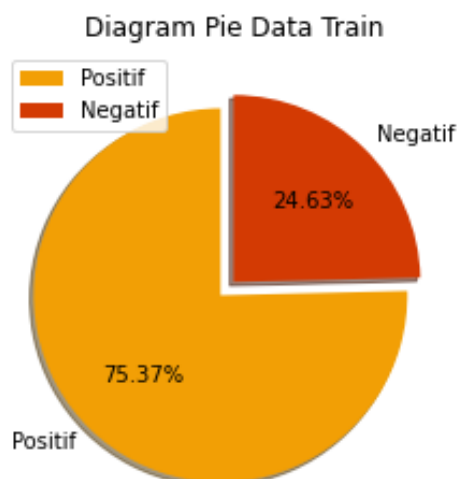
D_{ij} = elemen matriks data ke-ij

y_i = kelas atau label data ke-i

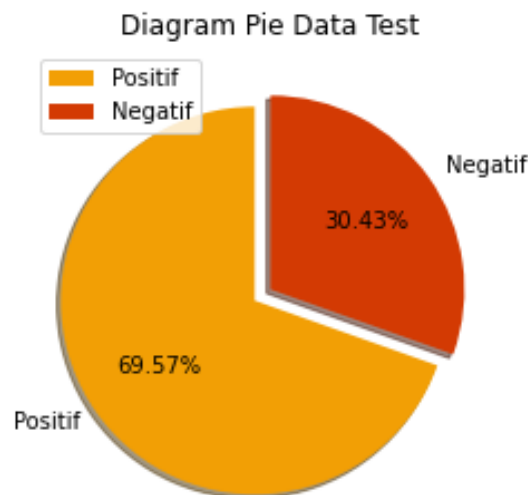
y_j = kelas atau label data ke-j

λ = turunan batas teoritis

Langkah awal dalam *Support Vector Machine* yaitu membagi data menjadi data *training* dan data *testing*. Pembagian data yaitu 90 % data *training* dan 10% data *testing* yang ditunjukkan pada gambar 8 dan gambar 9, dimana hasil pembagian tersebut menghasilkan data *train* sebanyak 306 positif dan 100 tweets negatif. Sedangkan hasil data test sebanyak 32 tweets positif dan 14 tweets negatif.

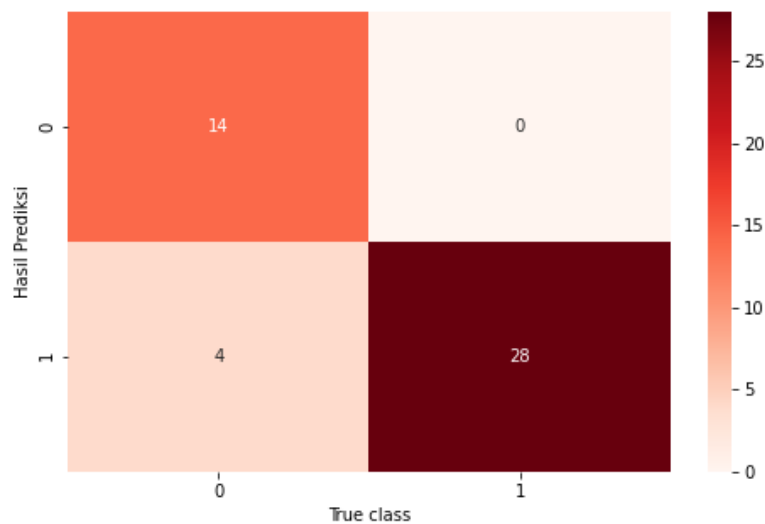


Gambar 8. Diagram pie hasil pembagian data train



Gambar 9. Diagram pie hasil pembagian data test

Berdasarkan hasil perhitungan *confusion matrix* pada algoritma *Support Vector Machine Kernel Linear* data tweets diatas prediksi benar sebanyak 42 tweets dan prediksi salah sebanyak 4 tweets. Hasil klasifikasi dengan algoritma SVM untuk data analisis sentimen pengguna IndiHome diperoleh nilai *accuracy* sebesar 91.3%. *Precession* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. *Precession* dengan menggunakan metode SVM sebesar 100%. *Recall* adalah rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. *Recall* dengan menggunakan metode SVM sebesar 88%. *F Measure* merupakan perbandingan rata-rata *precession* dan *recall* yang dibobotkan. *F Measure* pada metode SVM sebesar 93%. Adapun hasil *confusion matrix* ditunjukkan pada gambar 10.



Gambar 10. Hasil *confusion matrix*

Word cloud merupakan salah satu metode untuk menampilkan data *text* secara visual. *Word cloud* pada penelitian ini dibagi menjadi 2 yaitu *word cloud* sentimen positif dan *word cloud* sentimen negatif yang dapat dilihat pada gambar 11 dan gambar 12.



Gambar 11. *Word Cloud* sentimen positif



Gambar 12. *Word Cloud* sentimen negatif

Pada *word cloud* positif kata “*Virtual meet*” dan “semangat” berukuran paling besar dan menandakan bahwa kata tersebut banyak digunakan oleh pengguna. Dua kata tersebut berkaitan dengan event yang dilaksanakan oleh pihak IndiHome, hal tersebut menunjukkan bahwa pada sentimen positif pengguna Twitter sangat menyukai IndiHome jika melaksanakan sebuah event Pada *word cloud* negatif kata “*error*” dan “*lemot*” merupakan kata berukuran besar setelah kata “*indihom*” yang merupakan *keyword* pada saat *crawling* data. Hal ini merujuk pada tanggapan pengguna twitter yang ketika menggunakan layanan IndiHome sering mengalami *error* ataupun koneksi yang lemah.

4. KESIMPULAN

Sebelum dilakukan *clustering* pada *text tweets*, terlebih dahulu harus dilakukan *preprocessing* diantaranya : *text cleaning*, *case folding*, *stemming*, *stopword* dan *tokenizing*. Hasil ketepatan klasifikasi menggunakan *Support Vector Machine* dalam mengklasifikasi sentimen pengguna twitter mengenai layanan IndiHome diperoleh akurasi sebesar 91.3% Berdasarkan hasil tampilan *word cloud* pada sentimen positif kata “*virtual meet*”, “semangat”, “*meet asik*”, “*event*”. Sedangkan pada sentimen negatif kata “*error*”, “*lemot*”, “*ngelag*”, “*mola*” merupakan kata yang paling sering digunakan.

DAFTAR PUSTAKA

- [1] "Blog | IndiHome," *indihome.co.id*. <https://indihome.co.id> (diakses Jan 02, 2022).
- [2] "Sentiment Analysis and Opinion Mining | Synthesis Lectures on Human Language Technologies." <https://www.morganclaypool.com/doi/abs/10.2200/s00416ed1v01y201204hlt016> (diakses Des 15, 2021).
- [3] S. B. Bhonde dan J. R. Prasad, "Sentiment Analysis - Methods, Applications & Challenges," vol. 6, no. 6, hlm. 7, 2015.
- [4] R. Watrianthos, "Sentiment Analysis Of Traveloka App Using Naïve Bayes Classifier Method." INA-Rxiv, Jul 28, 2019. doi: 10.31227/osf.io/2dbe4.
- [5] C. B. Saputra, A. Muzakir, dan D. Udariansyah, "Analisis Sentimen Masyarakat Terhadap #2019GantiPresiden Berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier," *Bina Darma Conf. Comput. Sci. BDCCS*, vol. 1, no. 2, Art. no. 2, 2019.
- [6] B. W. Sari dan F. F. Haranto, "IMPLEMENTASI SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PELAYANAN TELKOM DAN BIZNET," *Pilar Nusa Mandiri J. Comput. Inf. Syst.*, vol. 15, no. 2, Art. no. 2, Sep 2019, doi: 10.33480/pilar.v15i2.699.
- [7] D. Darwis, E. S. Pratiwi, dan 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, Art. no. 1, Nov 2020, doi: 10.21107/edutic.v7i1.8779.
- [8] R. E. Husna, R. Wasono, dan M. A. Haris, "ANALISIS SENTIMEN PADA TWITTER MENGENAI NETFLIX DIBLOKIR TELKOM MENGGUNAKAN SUPPORT VECTOR MACHINE," *Pros. Semin. Nas. Venue Artik.-Ris. Inov. Reson.-Teori Dan Apl. Stat. VARIANSI*, vol. 2, no. 1, Art. no. 1, Mar 2021.
- [9] V. I. Santoso, G. Virginia, dan Y. Lukito, "PENERAPAN SENTIMENT ANALYSIS PADA HASIL EVALUASI DOSEN DENGAN METODE SUPPORT VECTOR MACHINE," *J. Transform.*, vol. 14, no. 2, Art. no. 2, Jan 2017, doi: 10.26623/transformatika.v14i2.439.
- [10] "Text Mining in Practice with R - Ted Kwartler - Google Buku." https://books.google.co.id/books?hl=id&lr=&id=2KxaDwAAQBAJ&oi=fnd&pg=PP11&dq=Text+Mining+in+Practice+with+R.+John+Wiley+%26+Sons+Ltd.&ots=_3iK9syfpW&sig=aSCCTea3CeIRN8CT2nI-671_xeM&redir_esc=y#v=onepage&q=Text%20Mining%20in%20Practice%20with%20R.%20John%20Wiley%20%26%20Sons%20Ltd.&f=false (diakses Des 15, 2021).
- [11] L. Kumar dan P. K. Bhatia, "TEXT MINING: CONCEPTS, PROCESS AND APPLICATIONS," hlm. 5, 2010.
- [12] H. L. Lukitowati dan L. Paryatno, "Evaluasi Pelaksanaan Webinar Pusdiklat Perdagangan Dengan Menggunakan Text Mining: Analisis Saran Peserta Webinar," *Cendekia Niaga*, vol. 5, no. 1, Art. no. 1, Jun 2021, doi: 10.52391/jcn.v5i1.574.
- [13] A. T. J. Harjanta, "Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining," *J. Inform. Upgris*, vol. 1, no. 1 Juni, Art. no. 1 Juni, 2015, doi: 10.26877/jiu.v1i1.