

Mengukur Kinerja Layanan Internet Indihome dari Opini Masyarakat Menggunakan Sentimen Analisis Twitter Dengan Metode Naïve Byes

Heribertus Himawan^{1*}, Rafif Murtadho², Dian Ferriswara³

^{1,2}Universitas Dian Nuswantoro

³Universitas Dr. Soetomo Surabaya

Email: ^{1*}himawan26@dsn.dinus.ac.id

ABSTRACT

In 2021, 77.02% of Indonesians said that their purpose for using the internet was to access social media. One of the social media used is Twitter. With many tweets that have been published through Twitter, these tweets can contain user opinions on a particular thing, it can be like an event in the surrounding Indihome. Through Twitter, users can discuss their complaints or satisfaction with the Indihome service. For that reason, a method is needed, namely sentiment analysis to find out whether the data contains negative or positive opinions. The author uses the Naïve Bayes method in conducting sentiment analysis on the opinions or opinions of Indihome service users on Twitter, to know how accurate the Naïve Bayes method is applied to sentiment analysis. After testing using the Naïve Bayes method, the results obtained are 82% accuracy, 78% precision, 84% recall, and 81% f1-score.

Keywords: Naïve Bayes, Sentiment analysis, Twitter, Indihome

ABSTRAK

Pada tahun 2021 jumlah masyarakat Indonesia yang mengakses internet ada sebanyak 77,02% dari jumlah penduduk Indonesia. Mayoritas pengguna internet Indonesia menggunakan internet untuk mengakses media sosial. Salah satu media sosial yang digunakan yaitu twitter. Dengan banyaknya tweet yang sudah ditulis melalui Twitter, tweet tersebut dapat berisi pendapat pengguna terhadap suatu hal tertentu, hal tersebut bisa seperti peristiwa masyarakat disekitar, seperti produk atau layanan. Contohnya yaitu perusahaan penyedia jasa internet (ISP), seperti Indihome. Melalui Twitter, pengguna bisa saling mendiskusikan keluhan atau kepuasan mereka kepada layanan Indihome. Untuk itu diperlukan suatu metode yaitu analisis sentimen untuk mengetahui apakah data tersebut mengandung pendapat negatif atau positif. Penulis menggunakan metode Naïve Bayes dalam melakukan analisis sentimen terhadap opini atau pendapat pengguna layanan Indihome di Twitter, dengan tujuan untuk mengetahui seberapa akurat metode Naïve Bayes yang diterapkan pada analisis sentimen. Setelah dilakukan pengujian menggunakan metode Naïve Bayes didapatkan hasil akurasi 82%, presisi 78%, recall 84%, dan f1-score 81%.

Kata kunci: Naïve Bayes, Analisis sentimen, Twitter, Indihome

A. Latar Belakang

Pada tahun 2021 jumlah penduduk Indonesia yang terkoneksi dengan internet ada sebanyak 210.026.769 juta atau 77,02% dari seluruh penduduk. Provider yang paling banyak digunakan oleh pengguna internet rumahan adalah Indihome, yaitu sebanyak 67,54% dari total pengguna internet (APJII, 2022).

Sebanyak 89,15% pengguna internet Indonesia menggunakan internet untuk mengakses media sosial dan salah satu media sosial yang populer digunakan adalah twitter.

Twitter merupakan media sosial yang dapat dipakai untuk menggali informasi tentang bisnis, hiburan, ekonomi, politik dan hal lainnya dengan

mudah sesuai dengan karakter pengguna media sosial yang kebanyakan hanya ingin berkomentar pendek untuk mengekspresikan sesuatu. Setiap kata atau kalimat yang dituliskan masyarakat memiliki makna yang tersirat, Dengan melakukan analisis yang disebut analisis sentimen maka perusahaan dapat mengetahui apakah komentar positif atau komentar negatif yang lebih banyak muncul terhadap hal-hal yang berhubungan dengan organisasi atau bisnisnya. Dengan memilah dan menjumlah *tweet* yang sudah ditulis melalui Twitter, maka pengguna dapat membuat kesimpulan tentang baik buruk, diterima atau ditolaknya suatu hal (Abdan, 2017). Hal yang dinilai tersebut bisa berupa peristiwa yang terjadi dimasyarakat disekitar, perilaku seseorang, ucapan seseorang atau tentang produk dan layanan.

Beberapa penelitian yang berhubungan dengan analisis sentiment twitter yang pernah dilakukan: Meningkatkan akurasi peramalan penjualan kendaraan bermotor dengan melakukan analisis sentiment media sosial Twitter dan nilai saham (Pai & Liu, 2018). Dalam penelitian menggunakan TSS (*Tweet Sentiment Score*) untuk melakukan prediksi tentang pasar keuangan, diperoleh akurasi 97,87% (Guo & Li, 2019). Wongkar dan Adresey melakukan analisis sentiment menggunakan algoritma Naïve Byes tentang kandidat presiden dan calon presiden Indonesia tahun 2019 dan mendapatkan hasil pasangan Joowi-Ma'ruf mempunyai sentimen positif 45,45% dan sentiment negatif 54,55%, sedangkan pasangan Prabowo-Sandy mendapatkan sentiment positif sebesar 44,32% dan sentiment negatif 55,68% dengan akurasi 73,34% (Wongkar &

Adresey, 2021). Sebagai salah satu platform media sosial terpopuler, perusahaan dapat menggunakan cuitan dalam Twitter sebagai media untuk mengetahui penerimaan atau kesan masyarakat pengguna terhadap sebuah produk yang telah diluncurkan, contohnya yang dilakukan oleh perusahaan penyedia jasa internet (ISP) Indihome.

Beberapa contoh cuitan masyarakat terkait layanan Indihome, diantaranya oleh akun kapitanindobn menyebutkan bahwa wifi yang digunakan oleh neneknya di kampung lebih baik daripada wifi Indihome yang ia digunakan di Jakarta. Akun openuptheskyee mengatakan internet Indihome yang digunakan selalu mengalami penurunan kecepatan setiap pagi. Akun oktamarimar mengatakan sudah 6 hari internet Indihome lambat setiap malam. Karena jumlah data tweet tersebut banyak, maka diperlukan metode untuk mengubah data tersebut menjadi sebuah informasi atau pengetahuan. Salah satu cara analisis data adalah analisis sentiment menggunakan teknik data mining atau khususnya adalah *text mining*.

B. Landasan Teoritis

Sentiment analysis atau dalam Bahasa Indonesianya analisis sentimen adalah sebuah teknik atau cara yang digunakan untuk mengidentifikasi bagaimana sebuah sentimen diekspresikan menggunakan teks dan bagaimana sentimen tersebut bisa dikategorikan sebagai sentimen positif maupun sentimen negatif (Nasukawa & Yi, 2003)

Text mining sendiri masih merupakan hal yang baru dalam riset ilmu komputer yang mencoba

memecahkan krisis kelebihan informasi dengan menggabungkan Teknik *data mining*, *machine learning*, *natural language processing*, *information retrieval* dan *knowledge management* (Feldman & Sanger, 2007).

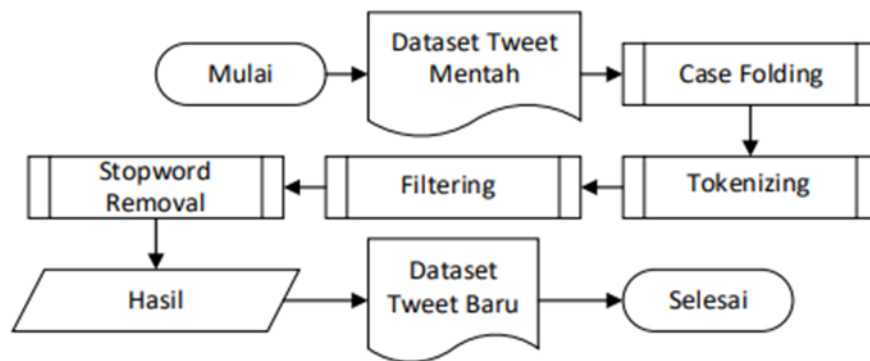
Jasa atau layanan menurut Lovelock, Wirtz, & Mussry (2010) didefinisikan sebagai kegiatan ekonomi yang ditawarkan kepada pihak lain dalam jangka waktu tertentu dan dikemas dalam sebuah bentuk kegiatan yang berdampak pada penerima, objek, maupun aset-aset lain yang menjadi tanggung jawab dari pembeli. Untuk mendapatkan keuntungan finansial dalam jangka panjang, perusahaan harus merancang dan memberikan layanan yang menyenangkan pelanggan, sehingga mereka memiliki pengalaman positif selama menikmati layanan tersebut (Lovelock, Patterson & Walker 2004).

Penelitian terkait analisis sentimen menggunakan algoritma Naïve Bayes pernah dilakukan oleh Abdan Syakuro (Harijiatno, 2017) yang mengelompokkan sentimen negatif dan positif pada sebuah media social dan mendapatkan hasil akurasi sebesar 88,8%. Penelitian serupa juga telah dilakukan oleh Servasius Dwi Harijiatno (2017), untuk mengetahui opini masyarakat terhadap tokoh publik di Indonesia dan mendapatkan nilai akurasi Naïve Bayes sebesar 72,941%. Analisis sentiment di Twitter dengan algoritma Naïve Bayes juga dilakukan oleh Bayhaqy dkk, untuk menganalisis kinerja sebuah e-commerce dan mendapatkan akurasi sebesar 77% (Bayhaqy et al., 2018). Dalam berbagai penelitian Naïve Bayes dinilai baik dalam melakukan analisis

sentiment terbukti dari nilai akurasi yang dinilai baik (Feldman & Sanger, 2007). Dalam penelitian mengenai Layanan Internet Banking menemukan bahwa layanan *cloud*, keamanan, *e-learning*, dan kualitas layanan merupakan 4 faktor yang secara signifikan menentukan kepuasan pelanggan (Feng Li et al., 2021). Penelitian ini akan melakukan analisis sentimen terhadap layanan internet Indihome menggunakan algoritma Naïve Bayes untuk melakukan klasifikasi dari cuitan-cuitan pengguna di media sosial Twitter.

C. Metode Penelitian

Data yang digunakan merupakan data primer yang dikumpulkan dengan cara melakukan *crawling tweet* berbahasa Indonesia dengan *keyword* “@indihome” melalui API Twitter dengan Tweepy, kemudian data *tweets* akan disimpan dalam format .csv. Selanjutnya akan dilakukan *Text Preprocessing* yang bertujuan untuk mengubah data tidak terstruktur menjadi data terstruktur dan dapat dikuantifikasi data sehingga menjadi nilai numerik (Feldman & Sanger, 2007). Adapun proses pengolahan data teks adalah melakukan *case folding* yaitu proses mengubah seluruh huruf menjadi huruf kecil, *tokenizing* adalah tahap pemotongan string masukan berdasarkan kata-kata yang menyusunnya atau dengan kata lain pemecahan kalimat menjadi kata, *filtering* yaitu adalah tahap mengambil kata penting dari hasil proses token, *stopword removal* yaitu proses membuang term yang tidak memiliki arti atau tidak relevan.



Gambar 1. Flowchart Proses *Text Preprocessing*

Pembobotan Kata TF-IDF

Term-weighting atau pembobotan kata merupakan proses memberikan bobot atau nilai term pada dokumen, yang nantinya akan digunakan untuk klasifikasi dokumen. TF = jumlah frekuensi kata terpilih / jumlah kata. IDF dihitung dengan rumus $IDF = \log(\text{jumlah dokumen} / \text{jumlah frekuensi kata terpilih})$. Hitung banyaknya dokumen yang memuat kata (DF) kemudian hitung IDF dengan rumus sebagai berikut :

$$IDF = \log \left(\frac{\text{Total Dokumen}}{\text{Dokumen Frekuensi}} \right) \quad (1)$$

Proses Klasifikasi Naïve Bayes

Selanjutnya dilakukan perhitungan probabilitas masing-masing kata diproses dengan menggunakan persamaan:

$$P(w|\text{positif/negatif}) = \frac{((nk|\text{positif/negatif})+1)}{((n,|\text{pos/neg})+|\text{kosakata})} \quad (2)$$

Keterangan :

$P(w|\text{positif/negatif})$ = Peluang kemunculan kata pada kategori.

w = Kata yang muncul pada sebuah kategori.

$((nk|\text{positif/negatif})+1)$ = Jumlah frekuensi kemunculan kata pada kategori.

nk = Kemunculan setiap kata pada kategori.

$|\text{kosakata}|$ = Jumlah semua kata dari semua kategori..

D. Hasil Penelitian dan Analisis

Persiapan Data

Pada penelitian ini digunakan data sebanyak 500 data tweet dimana data tersebut dibagi menjadi 250 tweet positif dan 250 tweet negatif. Lalu data tersebut akan dipisah untuk dijadikan data testing dan data training dengan perbandingan 20-80, dan didapatkan 100 data testing dan 400 data training.

Tabel 1 Data Tweet Indihome

Tweet
Semoga layanan indihome bisa lebih baik lagi @IndiHome
@indihome tlg internet lambat tidak bisa akses
Indihome lambat bgt. Semoga akses internet lebih baik lagi terimakasih

Text Preprocessing

Setelah data *tweet* didapatkan, data tersebut akan dilakukan proses *text preprocessing*. Adapun tahap preprocessing sebagai berikut:

- Case folding* adalah teknik merubah semua huruf yang ada dalam dokumen menjadi huruf kecil. Huruf

yang diubah hanya huruf yang ada pada alfabet dari ‘a’ sampai ‘z’. Contohnya: kalimat “Indihome lambat bgt. Semoga akses internet lebih baik lagi terimakasih” diubah menjadi “indihome lambat bgt. semoga akses internet lebih baik lagi terimakasih”.

- b. *Tokenizing* digunakan untuk memotong string masukan sesuai dengan kata penyusunnya. Atau bisa disebut sebagai teknik yang memisahkan setiap kata yang tersusun dari sebuah dokumen. Contoh: kalimat “indihome lambat bgt. semoga akses internet lebih baik lagi terimakasih” menghasilkan 9 token yaitu “indihome”, “lambat”, “semoga”, “akses”, “internet”, “lebih”, “baik”; “lagi”.
- c. *Filtering* yaitu teknik mengekstraksi

kata-kata yang dibutuhkan dari hasil *tokenizing*. Teknik *filtering* memiliki algoritma, yaitu *wordlist* dan *stoplist*. *Stoplist* digunakan untuk menghapus kata yang tidak penting, *wordlist* digunakan untuk menyimpan kata penting. *Stoplist* adalah kata-kata non-deskriptif, yang dapat dibuang menggunakan metode *bag-of-words*. Salah satu kata *stopwords* adalah “ke”, “aku”, “anda”, “adalah” dan seterusnya.

- d. *Stemming* merupakan teknik pencarian kata dasar yang sudah dilakukan saat proses *filtering*. Di tahap ini kata imbuhan pada suatu kata seperti kata “kerjaan” dimana “-an” adalah imbuhan akan dihapus supaya seluruh kata menjadi kata dasar.

Tabel 2 Hasil *Text Preprocessing*

Data Latih	Tweet	Kelas
1	semoga layanan indihome bisa lebih baik lagi	Positif
2	internet lambat tidak bisa akses	Negatif
Data Uji		
1	indihome lambat semoga akses internet lebih baik lagi terima kasih	?

Pembobotan Kata TF-IDF

Tabel 3 Hasil TF-IDF

No.	Kata	TF 1	TF 2	DF	IDF	TF-IDF	TF-IDF
1	semoga	1	0	1	0.301029996	0.301029996	
2	layanan	1	0	1	0.301029996	0.301029996	
3	indihome	1	0	1	0.301029996	0.301029996	
4	bisa	1	1	2	0	0	0
5	lebih	1	0	1	0.301029996	0.301029996	
6	baik	1	0	1	0.301029996	0.301029996	
7	lagi	1	0	1	0.301029996	0.301029996	
8	internet	0	1	1	0.301029996		0.301029996
9	lambat	0	1	1	0.301029996		0.301029996
10	tidak	0	1	1	0.301029996		0.301029996
11	akses	0	1	1	0.301029996		0.301029996

Klasiifikasi Naïve Bayes

Selanjutnya dilakukan perhitungan probabilitas masing-masing kata

diproses dengan menggunakan persamaan:

$$P(w|positif/negatif) = \frac{(nk|positif/negatif)+1}{(n,|pos/neg)+|kosakata|} \quad (3)$$

Keterangan :

- $P(w|positif/negatif)$ = Peluang kemunculan kata pada kategori.
 w = Kata yang muncul pada sebuah kategori.
 $(nk|positif/negatif) + 1$ = Jumlah frekuensi kemunculan kata pada kategori.
 nk = Kemunculan setiap kata pada kategori.
 $|kosakata|$ = Jumlah semua kata dari semua kategori.

Probabilitas kata “semoga”:

$$P("semoga"|"pos") = \frac{0.301029996 + 1}{5 + 11} = \mathbf{0.18131437475}$$

$$P("semoga"|"neg") = \frac{0 + 1}{3 + 11} = \mathbf{0.07142857142}$$

Probabilitas kata “indihome”:

$$P("indihome"|"pos") = \frac{0.301029996 + 1}{5 + 11} = \mathbf{0.18131437475}$$

$$P("indihome"|"neg") = \frac{0 + 1}{3 + 11} = \mathbf{0.07142857142}$$

Probabilitas kata “lebih”:

$$P("lebih"|"pos") = \frac{0.301029996 + 1}{5 + 11} = \mathbf{0.18131437475}$$

$$P("lebih"|"neg") = \frac{0 + 1}{3 + 11} = \mathbf{0.07142857142}$$

Probabilitas kata “baik”:

$$P("baik"|"pos") = \frac{0.301029996 + 1}{5 + 11} = \mathbf{0.18131437475}$$

$$P("baik"|"neg") = \frac{0 + 1}{3 + 11} = \mathbf{0.07142857142}$$

Probabilitas kata “lagi”:

$$P("lagi"|"pos") = \frac{0.301029996 + 1}{5 + 11} = \mathbf{0.18131437475}$$

$$P("lagi"|"neg") = \frac{0 + 1}{3 + 11} = \mathbf{0.07142857142}$$

Probabilitas kata “internet”:

$$P("internet"|"pos") = \frac{0 + 1}{5 + 11} = \mathbf{0.0625}$$

$$P("internet"|"neg") = \frac{0.301029996 + 1}{3 + 11} = \mathbf{0.092930714}$$

Probabilitas kata “lambat”:

$$P("lambat"|"pos") = \frac{0 + 1}{5 + 11} = \mathbf{0.0625}$$

$$P("lambat"|"neg") = \frac{0.301029996 + 1}{3 + 11} = \mathbf{0.092930714}$$

Probabilitas kata “akses”:

$$P(\text{"akses"}|\text{"pos"}) = \frac{0 + 1}{5 + 11} = \mathbf{0.0625}$$

$$P(\text{"akses"}|\text{"neg"}) = \frac{0.301029996 + 1}{3 + 11} = \mathbf{0.092930714}$$

Tahapan selanjutnya menghitung probabilitas kategori dengan rumus berikut:

$$P(\text{pos/neg}|d) = p(\text{pos/neg}) \times \prod p(a|\text{pos/neg}) \quad (4)$$

Keterangan:

$P(\text{pos/neg})$ = Peluang kemunculan kata pada kategori atau kelas dengan dokumen data uji.

d = Dokumen data baru.

$\prod p(a|\text{pos/neg})$ = Peluang kemunculan a_i pada kategori atau kelas.

a = Kata baru yang akan diuji.

Probabilitas kategori positif:

$$P(\text{uji}|\text{pos}) = p(\text{pos}) \times p(\text{semoga}|\text{pos}) \times p(\text{indihome}|\text{pos})$$

$$\times p(\text{lebih}|\text{pos}) \times p(\text{baik}|\text{pos}) \times p(\text{lagi}|\text{pos})$$

$$P(\text{uji}|\text{pos}) = 0,5 \times 0.18131437475 \times 0.18131437475$$

$$\times 0.18131437475 \times 0.18131437475 \times 0.18131437475$$

$$= \mathbf{0.45328593687}$$

$$P(\text{uji}|\text{pos}) = p(\text{pos}) \times p(\text{internet}|\text{pos}) \times p(\text{lambat}|\text{pos}) \times p(\text{akses}|\text{pos})$$

$$P(\text{uji}|\text{pos}) = 0,5 \times 0.0625 \times 0.0625 \times 0.0625$$

$$= \mathbf{0.09375}$$

Probabilitas kategori negatif:

$$P(\text{uji}|\text{neg}) = p(\text{neg}) \times p(\text{semoga}|\text{neg}) \times p(\text{indihome}|\text{neg})$$

$$\times p(\text{lebih}|\text{neg}) \times p(\text{baik}|\text{neg}) \times p(\text{lagi}|\text{neg})$$

$$P(\text{uji}|\text{neg}) = 0,5 \times 0.07142857142 \times 0.07142857142$$

$$\times 0.07142857142 \times 0.07142857142 \times 0.07142857142$$

$$= \mathbf{0.17857142855}$$

$$P(\text{uji}|\text{neg}) = p(\text{neg}) \times p(\text{internet}|\text{neg}) \times p(\text{lambat}|\text{neg}) \times p(\text{akses}|\text{neg})$$

$$P(\text{uji}|\text{neg}) = 0,5 \times 0.092930714 \times 0.092930714 \times 0.092930714$$

$$= \mathbf{0.139396071}$$

Nilai probabilitas tertinggi yaitu sebesar **0.45328593687** pada $p(\text{uji}|\text{pos})$, sehingga *tweet* tersebut diklasifikasikan kedalam kelas “**Positif**”.

Dari 500 data *tweet* didapatkan hasil perhitungan nilai akurasi sebesar 82%, akurasi antara hasil prediksi dengan data aktual (presisi) sebesar 78%, dan

tingkat keberhasilan sistem memprediksi suatu data (*recall*) sebesar 84%, dimana rata-rata nilai perbandingan tingkat presisi dan tingkat *recall* (*f1_score*) adalah 81%. Ada 39 *True Positive* (TP), 7 *False Positive* (FP), 43 *True Negative* (TN), dan 11 *False Negative* (FN).

```
MultinomialNB Accuracy: 0.82
MultinomialNB Precision: 0.78
MultinomialNB Recall: 0.8478260869565217
MultinomialNB f1_score: 0.8125
confusion matrix:
[[39  7]
 [11 43]]
=====
              precision    recall  f1-score   support

negatif         0.78         0.85         0.81         46
positif         0.86         0.80         0.83         54

accuracy                0.82
macro avg              0.82         0.82         0.82         100
weighted avg          0.82         0.82         0.82         100
```

Gambar 1 Hasil Akurasi Naive Bayes

E. Kesimpulan dan Rekomendasi

Hasil analisis sentiment terhadap kualitas layanan Indihome sebagai penyedia jaringan internet menunjukkan bahwa sentimen negatif masih lebih besar (43%) dari sentimen positif (39%). Dari hasil ini diharapkan Indihome melakukan peningkatan kinerja karena masih lebih banyak sentimen negative disbanding sentiment positifnya. Berdasarkan pembahasan juga didapatkan simpulan, bahwa Algoritma Naive Bayes dapat digunakan untuk menganalisis sentiment dari data yang diambil dari Twitter (cuitan Twitter) dengan hasil akurasi hingga 82%.

REFERENSI

A. Bayhaqy, S. Sfenrianto, K. Nainggolan and E. R. Kaburuan (2018), "Sentiment Analysis about E-Commerce from Tweets Using Decision Tree, K-Nearest Neighbor, and Naive Bayes," 2018 International Conference on Orange Technologies (ICOT), 2018, pp. 1-6, doi: 10.1109/ICOT.2018.8705796.

Abdan S., (2017), Analisis Sentimen Masyarakat Terhadap E-Commerce Pada Media Sosial Menggunakan Metode Naive Bayes Classifier

(NBC) Dengan Seleksi Fitur Information Gain (IG), *Tesis*, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim, Malang.

APJII, 2020, *Laporan Survei Internet APJII 2019 - 2020 [Q2]*, APJII, Jakarta.

Darma, dan I. M. B. S., (2017), Penerapan Sentimen Analisis Acara Televisi Pada Twitter Menggunakan Support Vector Machine Dan Algoritma Genetika Sebagai Metode Seleksi Fitur, *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, No.43, Vol.2, 998-1007.

Feldman, R. dan James, S. (2007). *The Text Mining Handbook*. New York: Cambridge.

Feng Li, Hui Lu, Meiqian Hou, Kangle Cui, Mehdi Darbandi (2021), Customer satisfaction with bank services: The role of cloud services, security, e-learning and service quality, *Technology in Society*, Volume 64, 2021, 101487, ISSN 0160-791X

Guo, X., & Li, J. (2019). A novel twitter sentiment analysis model with baseline correlation for financial market prediction with improved efficiency. In *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)* (pp. 472-477). IEEE.

Harijati, Servasius Dwi (2019) Analisis sentimen pada twitter menggunakan Multinomial Naive Bayes. Skripsi thesis, Sanata Dharma University.

Lovelock, C, Patterson, P & Walker, R 2004, *Services Marketing*, Prentice Hall Australia Pty Ltd., Riverwood, NSW.

- Lovelock, Christopher., Jochen, Wirtz., & Jacky, Mussry. (2010). *Pemasaran Jasa – Perspektif Indonesia Jilid 1*. Jakarta: Penerbit Erlangga.
- M. Wongkar and A. Angdresey (2019), "Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler: Twitter," 2019 Fourth International Conference on Informatics and Computing (ICIC), 2019, pp. 1-5, doi: 10.1109/ICIC47613.2019.8985884.
- Nasukawa, T. & Yi, J., (2003). Sentiment Analysis: Capturing Favorability Using Natural Language Processing. In *Proceedings of the 2nd International Conference on Knowledge Capture*. pp. 70–77
- Pai, P. F., & Liu, C. H. (2018). Predicting vehicle sales by sentiment analysis of Twitter data and stock market values. *IEEE Access*, 6, 57655-57662.