

PENGARUH SENTIMEN DI SOSIAL MEDIA DENGAN HARGA SAHAM PERUSAHAAN

Musthofa Galih Pradana¹, Azriel Christian Nurcahyo², Pujo Hari Saputro³

^{1,3}Universitas Alma Ata
Yogyakarta, Indonesia

²STIM Shanti Bhuana
Bengkayang, Kalimantan Barat

¹mgalihpradana@almaata.ac.id, ²azriel@shantibhuana.ac.id, ³pujo@almaata.ac.id

Abstrak

Sentimen adalah proses komputasional dalam mengidentifikasi dan mengategorikan opini-opini dalam bentuk potongan teks, khususnya untuk mengukur maksud si pembuat potongan teks terhadap topik tertentu, dapat bernilai positif, negatif, atau netral. Dalam konteks layanan perusahaan, sentimen yang sering muncul biasanya adalah sentimen yang bernilai positif dalam bentuk pujian dan apresiasi maupun sentimen bernilai negatif dalam bentuk complain. Ketika complain dilakukan oleh pelanggan, maka pihak perusahaan harus dapat melakukan tindakan untuk menanggulangi complain tersebut, yang dimungkinkan dapat berimbas terhadap kredibilitas perusahaan, bahkan dapat berimbas pada harga saham. Sosial media menjadi tempat yang dapat digunakan untuk menyampaikan keluhan maupun review positif dari pelanggan terhadap layanan yang diberikan oleh perusahaan. Proses analisis sentimen pelanggan terhadap layanan perusahaan perlu dilakukan untuk mengetahui seberapa besar jumlah sentimen yang muncul menggunakan metode KNN dengan studi kasus pada perusahaan telekomunikasi di Indonesia. Kemudian dilakukan analisis regresi untuk menilai apakah jumlah sentimen pelanggan berpengaruh terhadap harga saham perusahaan. Penelitian ini menghasilkan sentimen positif sebesar 16% , sentimen negatif sebesar 78 % dan sentimen netral sebesar 6%. Metode KNN menghasilkan akurasi paling tinggi sebesar 79,06% pada cross validation = 4. Jumlah sentimen negatif dan harga saham memiliki nilai regresi dan nilai korelasi sebesar 0,46 atau dengan kata lain tidak memiliki keterkaitan satu sama lain. Sentimen positif menghasilkan nilai regresi dan korelasi sebesar 0,02 dengan artian memiliki keterkaitan satu sama lain.

Kata Kunci: *Sentimen, KNN, Regresi, Cross Validation.*

Abstract

Sentiment is a computational process in identifying and categorizing opinions in the form of pieces of text, specifically to measure the intent of the creator of the piece of text to a certain topic, can be positive, negative, or neutral. In the context of corporate services, the sentiment that often arises is usually positive sentiment in the form of praise and appreciation and negative sentiment in the form of complaints. When a complaint is made by a customer, then the company must be able to take action to overcome the complaint, which is possible to have an impact on the company's credibility, and even can affect the stock price. Social media is a place that can be used to submit complaints and positive reviews from customers about the services provided by the company. The process of analyzing customer sentiments for company services needs to be done to find out how much sentiment appears using the KNN method with a case study on telecommunications companies in Indonesia. Then a regression analysis is performed to assess whether the amount of customer sentiment affects the company's stock price. This research resulted in positive sentiment of 16%, negative sentiment of 78% and neutral sentiment of 6%. The KNN method produces the highest accuracy of 79.06% at cross validation = 4. The number of negative sentiments and share prices have a regression value and a correlation value of 0.46 or in other words do not have an attachment to each other. Positive sentiments produce a regression value and correlation of 0.02 in the sense of having an attachment to each other.

Keywords: *Sentiment, KNN, Regression, Cross Validation.*

PENDAHULUAN

Sentimen Negatif merupakan opini yang dikeluarkan di media sosial yang memiliki kesan kurang baik bagi pihak yang disebut, sebaliknya sentimen positif menimbulkan kesan yang baik bagi pihak yang terlibat atau yang disebutkan. Menurut Lovelock & Wright (2002) komplain adalah bentuk ekspresi formal tentang ketidaksukaan atau ketidakpuasan terhadap beberapa aspek yang diterima seseorang. Komplain sangat wajar diterima oleh sebuah perusahaan baik dari aspek produk maupun layanan yang diberikan belum mampu memuaskan keinginan pelanggan. Perusahaan harus siap dengan segala bentuk komplain yang bias saja diterima. Di era media sosial yang semakin maju, semakin canggih menyampaikan keluhan menjadi sangat mudah, pelanggan tinggal mengakses media sosial dan berharap keluhan yang dialami segera ditindaklanjuti dengan cepat, dibandingkan melayangkan komplain secara konvensional.

Perusahaan harus memberikan pelayanan semaksimal mungkin untuk meminimalisir adanya keluhan pelanggan. Keluhan yang disampaikan oleh pelanggan di media sosial bisa saja dapat membentuk stigma buruk untuk perusahaan, karena sifat media sosial yang tidak ada sekat membuat siapapun dapat melihat komplain yang dilakukan oleh pelanggan lainnya. Hal ini dapat berimbas pada kredibilitas perusahaan, bahkan bisa saja calon pelanggan memutuskan untuk tidak jadi berlangganan atau menggunakan jasa dan produk dari perusahaan terkait akibat testimoni yang buruk dari pelanggan lama. Perusahaan tidak boleh mengesampingkan sentimen yang timbul ini. Lebih jauh lagi,

komplain juga kemungkinan besar dapat mempengaruhi nilai sebuah perusahaan atau nilai saham yang anjlok seiring banyaknya komplain yang diterima. Sebaliknya sentimen positif pelanggan di sosial media berpeluang memiliki keterikatan terhadap harga saham, karena bukan tidak mungkin bahwa review positif dari seseorang akan menimbulkan keyakinan orang lain yang hendak melakukan mitra bisnis dengan perusahaan.

Berdasarkan permasalahan tersebut, diperlukan sentimen analisis untuk mendapatkan informasi apakah harga saham akan terpengaruh dengan sentimen pada sebuah perusahaan. Sumber data yang akan dijadikan acuan adalah dari sosial media Facebook Perusahaan telekomunikasi di Indonesia dengan menggunakan metode KNN.

Tinjauan Pustaka

Penelitian tentang sentimen analisis pernah dilakukan oleh Ghaith Abdulsattar dkk penelitian ini bertujuan dalam membuat framework dalam sentimen analisis yang mengklasifikasikan sentimen ke dalam 3 jenis negatif, positif, dan netral. Proses pelabelan dilakukan secara otomatis. Hasilnya dapat mengklasifikasikan harga di pasaran untuk dijadikan bahan pertimbangan pihak terkait seperti investor, perusahaan, dan para peneliti (Abdulsattar dkk., 2018).

Penelitian kedua dibuat oleh Yefeng Ruan dkk, penelitian ini melakukan sentimen di media sosial twitter dan melakukan analisis menggunakan korelasi pearson, dan linier regresion untuk menghasilkan analisis harga pasar (Ruan dkk., 2018). Penelitian tentang data mining juga pernah dituliskan kaharudin dkk, yang

membandingkan metode KNN dan Naive Bayes. Kesimpulan dari penelitian ini adalah metode yang menghasilkan akurasi yang lebih baik adalah metode K-NearestNeighbor dengan nilai K=5 dengan tingkat akurasi 78% (Kaharudin, Musthofa Galih Pradana, 2019).

Musthofa Galih Pradana pernah menuliskan penggunaan Term Document Matrix dan Wordcloud pada R Studio dalam penelitian text mining. Penelitian menyimpulkan bahwa penerapan *term document matrix* dan *word cloud* saling melengkapi dalam proses text mining, dan memudahkan dalam pengolahan informasi berbasis teks(Pradana, 2020). Penelitian selanjutnya dari Rohan Pimprikar dkk, penelitian ini melakukan prediksi harga menggunakan sentimen analisis metode *Support Vector Machine* dan *Linier Regression*. Metode SVM menghasilkan akurasi sebesar 60% dan Linier Regresion menghasilkan 82% akurasi (Rohan Pimprikar1, S. Ramachandran2, 2017).

Ayman E. Khedr menulis tentang cara memprediksi harga di pasaran menggunakan teknik data mining dan sentimen analisis. Penelitian ini menghasilkan akurasi dari 72.73% sampai 86.21% menggunakan metode Naïve Bayes, sedangkan pada percobaan kedua dapat menghasilkan akurasi sebesar 89.80% (Khedr, 2017). Shri Bharathi dan Angelina Geetha melakukan penelitian tentang prediksi harga di pasaran agar lebih efektif, penelitian in berhasil meningkatkan 14.43% akurasi prediksi, ketika dibandingkan dengan metode ID3 dan C4.5(Bharathi, 2017).

Berikutnya penelitian dari Nuno Oliveira dkk yang melakukan prediksi harga pasar menggunakan dataset Twitter,

penelitian ini berhasil melakukan prediksi returns, volatility, trading volume and survey sentiment indices yang berguna dalam stock market (Oliveira et al., 2017). Penelitian ke 7 dari Abhishek Kumar dkk, yang melakukan analisis sentimen pada financial text, penelitian ini membangun 2 system, yang pertama yakni Support Vector Classifier Logistic Regression, kemudian yang kedua adalah Support Vector Regression menggunakan word embeddings, lexicon features, dan nilai PMI sebagai fitur (Kumar et al., 2017).

Sara Rosenthal dkk melakukan penelitian tentang sentimen analisis dan berhasil memperkenalkan bahasa arabic dalam analisisnya. Klasifikasi tweet dibagi menjadi 2 poin dan 5 poin ordinal scale(Rosenthal et al., 2017). Penelitian Sentimen Analisis juga dilakukan María del Pilar Salas-Zárate dkk yang berhasil mendapatkan kesimpulan akurasi precision 81.93%, recall 81.13%, dan *F-measure of* 81.24% untuk tema diabetes(Salas-zárate et al., 2017).

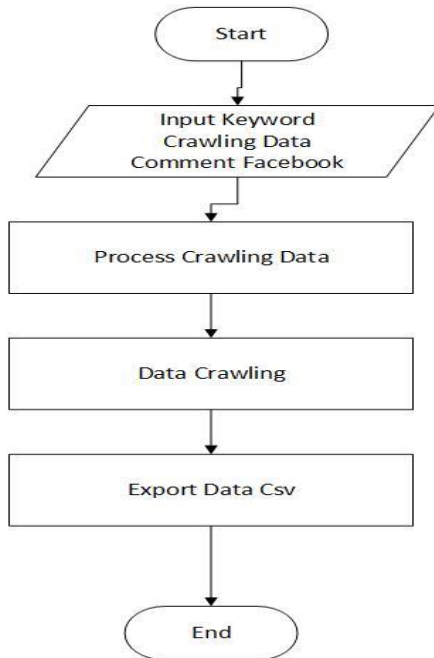
Penelitian rujukan terakhir dari Munir Ahmad dkk, penelitian ini menerapkan metode SVM di dalam aplikasi weka yang menghasilkan kesimpulan pada 2 dataset yang digunakan menghasilkan akurasi Self-Driving Cars sebesar 59.91% dan dataset Apple sebesar 71.2%(Ahmad & Ali, 2017).

METODE PENELITIAN

Metode Pengumpulan Data

Pengumpulan data dilakukan melalui proses crawling. Proses pengambilan data didapatkan dari sumber sosial media Facebook. Data yang diambil dari Facebook, masih *unstructured*, masih

banyak sampah yang perlu diolah di proses pengolahan data. Proses pengumpulan data ditunjukkan pada Gambar 1.



Gambar 1. Alur Pengumpulan Data

Metode Pengolahan Data

Data hasil *crawling* diolah dan dipilih yang sesuai dengan kebutuhan analisis. Tahap yang dilakukan pada pengolahan data adalah *preprocessing*, adapun langkahnya adalah :

- a. *Filtering duplicate*
- b. *Case folding*
- c. *Cleaning*
- d. *Tokenizing / parsing*
- e. *Filtering*

Implementasi

Setelah melalui tahapan pengolahan data dengan *preprocessing*, maka akan dilakukan proses implementasi ke dalam metode KNN dan di cari hubungannya dengan harga saham menggunakan analisis regresi. Selain itu juga akan dicari kata apa

yang sering muncul dalam sentimen pelanggan di sosial media.

Kesimpulan

Tahap yang terakhir adalah penarikan kesimpulan dari apa yang sudah dikerjakan dari tahap pertama sampai ketiga, sehingga di dapatkan informasi apakah sentiment berpengaruh terhadap harga saham.

HASIL PENELITIAN DAN PEMBAHASAN

Dataset

Dataset yang digunakan dalam penelitian adalah data hasil *crawling* data di Sosial Media Facebook pada periode Januari 2017 sampai Desember 2018 dengan jumlah data mentah sebanyak 2.850. Setelah data didapatkan dari facebook. Proses setelah mendapatkan data adalah dengan mengolah dan membersihkan data tersebut sehingga dapat menjadi informasi yang berguna. Dari data awal sejumlah 2850 setelah dibersihkan menjadi 1605 yang selanjutnya dilakukan pemilihan data yang memiliki sentimen terhadap perusahaan. Data detail ditunjukkan pada Tabel 1.

Tabel 1. Dataset

<i>Periode Waktu</i>	<i>Jumlah Comment</i>
Januari 2017	257
Februari 2017	13
Maret 2017	63
April 2017	14
Juli 2017	38

Agustus 2017	45
September 2017	150
Oktober 2017	316
November 2017	57
Desember 2017	161
Januari 2018	105
Februari 2018	163
Maret 2018	15
April 2018	45
Mei 2018	10
Juni 2018	36
Juli 2018	37
Agustus 2018	57
September 2018	19
Desember 2018	2
Total	1605

2017			
April 2017	14	0	0
Juli 2017	14	21	3
Agustus 2017	0	43	2
September 2017	19	123	8
Oktober 2017	85	209	22
November 2017	0	52	5
Desember 2017	30	115	16
Januari 2018	5	88	12
Februari 2018	19	135	9
Maret 2018	2	13	0
April 2018	4	39	2
Mei 2018	4	6	0
Juni 2018	7	26	3
Juli 2018	1	35	1
Agustus 2018	3	46	8
September 2018	6	8	5
Desember 2018	1	1	0

Pengolahan Data

Data setelah proses pembersihan data, selanjutnya dilakukan pelabelan sentimen yang dibagi menjadi 3 jenis yaitu sentimen positif, negatif dan netral. Data pengolahan data ditunjukkan pada Tabel 2.

Tabel 2. Klasifikasi Sentimen

<i>Periode Waktu</i>	<i>Positif</i>	<i>Negatif</i>	<i>Netral</i>
Januari 2017	29	227	1
Februari 2017	1	12	0
Maret	6	56	1

Harga Saham

Harga saham digunakan sebagai variabel yang akan dicari hubungan dengan jumlah sentimen di *comment* facebook. Data harga saham yang digunakan adalah periode waktu yang sama dengan periode waktu

crawling data. Data harga saham ditunjukkan pada Tabel 3.

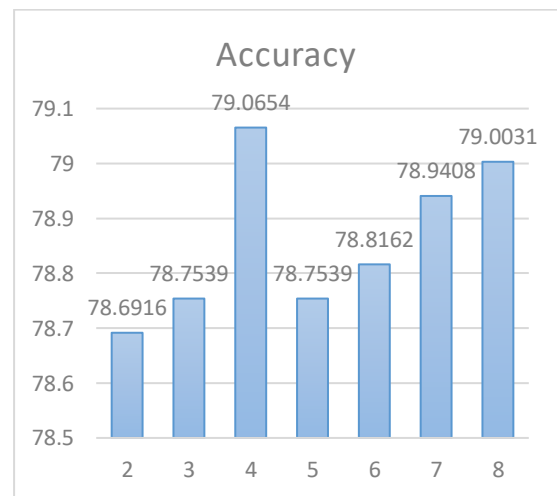
Tabel 3. Harga Saham

<i>Periode</i>	<i>Harga Saham</i>	<i>Keterangan</i>
Januari 2017	1,28B	-
Februari 2017	1,49B	Naik
Maret 2017	1,79B	Naik
April 2017	1,58B	Turun
Juli 2017	1,68B	Naik
Agustus 2017	1,25B	Turun
September 2017	1,39B	Naik
Oktober 2017	3,76B	Naik
November 2017	2,34B	Turun
Desember 2017	1,29B	Turun
Januari 2018	583,32M	Turun
Februari 2018	378,00M	Turun
Maret 2018	401,99M	Naik
April 2018	386,79M	Turun
Mei 2018	475,31M	Naik
Juni 2018	849,81M	Naik
Juli 2018	974,13M	Naik
Agustus 2018	684,02M	Turun

September 2018	295,33M	Turun
Desember 2018	462,51M	Naik

Metode KNN

Metode KNN yang diterapkan untuk mengklasifikasikan data diuji dengan teknik *cross validation*. Pengujian *cross validation* merupakan proses membagi data secara acak ke dalam beberapa bagian. Proses pengujian dimulai dengan pembentukan model dengan data pada bagian pertama. Model yang terbentuk akan diujikan pada bagian data yang tersisa. Setelah itu proses akurasi dihitung dengan melihat seberapa banyak data yang sudah terklasifikasi dengan benar. Data hasil pengujian ditunjukkan Gambar 2.



Gambar 2. Akurasi Cross Validation

Berdasarkan pengujian cross validation metode KNN, didapatkan nilai tertinggi pada nilai cross validation = 4 atau dengan nilai 79,06 %. Adapun detail pada nilai *cross validation* = 4 ditunjukkan pada Tabel 4.

Tabel 4. Detail Akurasi

<i>Sentimen</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
Positif	0,769	0,080	0,144
Negatif	0,792	0,994	0,881
Netral	0,000	0,000	0,000

Setelah diketahui besaran sentiment analisis pada proses pengolahan data, maka akan dicari apakah pada periode tertentu akan berpengaruh dengan harga saham menggunakan analisis regresi. Analisis regresi yang dilakukan dibagi menjadi 2 yaitu regresi terhadap sentimen positif dan regresi terhadap sentiment negatif.

Ketentuan dalam melakukan analisis regresi sebagai berikut :

Cara membaca korelasi adalah dengan:

Berdasarkan arahnya :

Apabila positif (+) maka hubungannya searah, artinya jika 1 variabel naik maka yang lain ikut naik

Apabila negatif (-) maka hubungannya berlawanan, artinya jika 1 variabel naik maka yang lain turun.

Berdasarkan nilai derajat korelasinya baik positif maupun negatif :

0,7 s/d 1 = kuat

0,4 s/d 0,7 = sedang

0,2 s/d 0,4 = rendah

< 0,2 = lemah

Cara melihat nilai korelasi kedua adalah dengan melihat jika angka probabilitas > 0,05 maka tidak ada hubungan yang signifikan antara kedua variable tersebut dan sebaliknya.

Cara melakukan analisis regresi diperlukan hipotesis sebagai berikut:

H0: Tidak ada hubungan linier antara variabel jumlah sentiment dengan harga saham

H1: Ada hubungan linier antara variable jumlah sentiment dengan harga saham

Pengujian dilakukan dengan menggunakan Sig dengan ketentuan:

Jika angka signifikansi < 0,05; H0 ditolak dan H1 diterima.

Jika angka signifikansi > 0,05; H0 diterima dan H1 ditolak.

Sentimen Negatif

Nilai korelasi 0,450. Artinya, hubungan kedua variable tersebut sedang. Korelasi positif menunjukkan bahwa hubungan antara jumlah sentiment dengan harga saham searah. Dalam arti jika jumlah sentimen negatif semakin sering maka harga saham akan menurun. Nilai korelasi sebesar 0,46 > 0,05 = Tidak ada hubungan yang signifikan antar variable. Gambar hasil analisis korelasi ditunjukkan pada Gambar 3.

Correlations

		Jumlah Sentimen	Harga Saham
Jumlah Sentimen	Pearson Correlation	1	,450*
	Sig. (2-tailed)		,046
	N	20	20
Harga Saham	Pearson Correlation	,450*	1
	Sig. (2-tailed)	,046	
	N	20	20

*. Correlation is significant at the 0.05 level (2-tailed).

Gambar 3. Korelasi Sentimen Negatif

Besarnya pengaruh variable Jumlah Sentimen terhadap Harga Saham adalah sebesar 20,3%. Hasil analisis ditunjukkan pada Gambar 4.

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,450 ^a	,203	,159	967161513 20,033

a. Predictors: (Constant), Jumlah Sentimen

b. Dependent Variable: Harga Saham

Gambar 4. Analisis Model Summary

Perhitungan angka signifikansi $0,46 > 0,05$; H_0 diterima dan H_1 ditolak. Artinya, tidak ada hubungan linier antara kedua variabel, maka sentiment negative tidak mempengaruhi harga saham. Hasil analisis anova ditunjukkan pada Gambar 5.

ANOVA^a

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	4287132 7990501 5150000 00,000	1	4287132 7990501 5150000 00,000	4,583	,046 ^b
Residual	1683722 5067087 1600000 000,000	18	9354013 9261595 3400000 0,000		
Total	2112435 7866137 3120000 000,000	19			

a. Dependent Variable: Harga Saham

b. Predictors: (Constant), Jumlah Sentimen³

Gambar 5. Analisis Anova Regresi

Sentimen Positif

Nilai korelasi 0,659. Artinya, hubungan kedua variable tersebut sedang. Korelasi positif menunjukkan bahwa hubungan antara jumlah sentiment dengan harga saham searah. Dalam arti jika jumlah sentiment negative semakin sering maka harga saham akan menurun. Nilai korelasi sebesar $0,02 < 0,05 =$ Ada hubungan yang signifikan antar variable. Gambar hasil analisis korelasi ditunjukkan pada Gambar 6.

Correlations

		Sentimen Positif	Harga Saham
Sentimen Positif	Pearson Correlation	1	,659**
	Sig. (2-tailed)		,002
	N	20	20
Harga Saham	Pearson Correlation	,659**	1
	Sig. (2-tailed)	,002	
	N	20	20

** . Correlation is significant at the 0.01 level (2-tailed).

Gambar 6. Korelasi Sentimen Positif

Besarnya pengaruh variable Jumlah Sentimen terhadap Harga Saham adalah sebesar 43,4%. Hasil analisis ditunjukkan pada Gambar 7.

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,659 ^a	,434	,403	8146870384 3,636

a. Predictors: (Constant), Sentimen Positif

b. Dependent Variable: Harga Saham

Gambar 7. Analisis Model Summary

Perhitungan angka signifikansi $0,02 < 0,05$; H_0 ditolak dan H_1 diterima. Artinya ada hubungan linier antara kedua variabel, maka sentiment positif mempengaruhi harga saham. Hasil analisis anova ditunjukkan pada Gambar 8.

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	917748 839540 544100 00000,0 00	1	917748 839540 544100 00000,0 00	13,8 27	,002 b
Residual	119468 694707 318710 000000, 000	18	663714 970596 215000 0000,00 0		
Total	211243 578661 373120 000000, 000	19			

a. Dependent Variable: Harga Saham

b. Predictors: (Constant), Sentimen Positif

Gambar 8. Analisis Anova Regresi

Wordcloud

Wordcloud merupakan fitur yang dapat digunakan untuk mengidentifikasi kata yang paling sering atau paling banyak digunakan pengguna sosial media saat berinteraksi dengan perusahaan. Pada penelitian ini, sebenarnya kata dengan frekuensi paling tinggi adalah kata nama perusahaan, namun penulis hilangkan untuk menghilangkan personalisasi, beserta kata yang berhubungan dengan produk dan layanannya. Adapun setelah

pemilihan kata yang dihilangkan tersebut data frekuensi pada wordcloud ditunjukkan pada Tabel 5.

Tabel 5. Hasil Wordcloud

Kata	Frekuensi
Indonesia	306
Internet	235
Hari	223
Tolong	205
Gangguan	190
Jaringan	187
Bayar	177
Mati	175
Sampai	137
Pelanggan	128

Berdasarkan 10 terbesar kata yang sering digunakan pengguna sosial media, data lebih detail atau lebih banyak tentang kata pada wordcloud ditunjukkan pada Gambar 9



Gambar 9. Wordcloud 100 Kata

KESIMPULAN DAN SARAN

Berdasarkan penelitian yang telah dilakukan dapat ditarik kesimpulan sebagai berikut :

1. Sentimen Positif sebesar 16 % , sentimen negatif 78 % dan sentimen netral sebesar 6%
2. Sentimen Negatif tidak memiliki hubungan dengan harga saham.
3. Sentimen Positif memiliki hubungan dengan harga saham.
4. Kata yang paling banyak digunakan pengguna adalah Indonesia dengan jumlah sebesar 306 kata.

Adapun saran dalam penelitian ini adalah data dapat diperkaya dari sumber data lainnya seperti Twitter, Youtube, Instagram dan Sosial Media lainnya agar lebih mendapatkan tingkat kepastian yang lebih baik.

DAFTAR PUSTAKA

- Abdulsattar, G., Alkubaisi, A. J., Kamaruddin, S. S., & Husni, H. (2018). *Conceptual Framework for Stock Market Classification Model Using Sentiment Analysis on Twitter Based on Hybrid Naïve Bayes Classifiers*. 7, 57–61.
- Ahmad, M., & Ali, I. (2017). *Sentiment Analysis of Tweets using SVM*. 177(5), 25–29.
- Bharathi, S. (2017). *Sentiment Analysis for Effective Stock Market Prediction*. 10(3), 146–154. <https://doi.org/10.22266/ijies2017.063016>
- Kaharudin, K., Pradana, M. G., & Kusriani, K. (2019). PREDIKSI CUSTOMER CHURN PERUSAHAAN

TELEKOMUNIKASI

MENGGUNAKAN NAÏVE BAYES DAN K-NEAREST NEIGHBOR. *Informasi Interaktif*, 4(3), 165-171.

- Khedr, A. E. (2017). *Predicting Stock Market Behavior using Data Mining Technique and News Sentiment Analysis*. July, 22–30. <https://doi.org/10.5815/ijisa.2017.07.03>
- Kumar, A., Sethi, A., Akhtar, S., & Ekbal, A. (2017). *Sentiment Prediction in Financial Text*.
- Oliveira, N., Cortez, P., & Areal, N. (2017). The impact of microblogging data for stock market prediction: Using Twitter to predict returns , volatility , trading volume and survey sentiment indices. *Expert Systems With Applications*, 73, 125–144. <https://doi.org/10.1016/j.eswa.2016.12.036>
- Pradana, M. G. (2020). *PENGGUNAAN FITUR WORDCLOUD DAN DOCUMENT TERM MATRIX DALAM TEXT MINING*.
- Rohan Pimprikar¹, S. Ramachandran², K. S. (2017). *USE OF MACHINE LEARNING ALGORITHMS AND TWITTER SENTIMENT ANALYSIS FOR STOCK MARKET PREDICTION*. 115(6), 521–526.
- Rosenthal, S., Farra, N., & Nakov, P. (2017). *Sentiment Analysis in Twitter*. 502–518.
- Ruan, Y., Duresi, A., & Alfantoukh, L. (2018). Using Twitter trust network for stock market analysis. *Knowledge-Based Systems*, 0, 1–12. <https://doi.org/10.1016/j.knosys.2018.01.016>
- Salas-zárate, M. P., Medina-moreira, J.,

Lagos-ortiz, K., Luna-aveiga, H., Rodríguez-garcía, M. Á., & Valencia-garcía, R. (2017). *Sentiment Analysis on Tweets about Diabetes: An Aspect-Level Approach*. 2017.