

TUGAS AKHIR

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP
INVESTASI KEUANGAN DI INDONESIA MENGGUNAKAN METODE
NAIVE BAYES**



Oleh:

Andreas Danny Agus W

1822250060

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER DAN REKAYASA
UNIVERSITAS MULTI DATA PALEMBANG
PALEMBANG
2023**

**Fakultas Ilmu Komputer dan Rekayasa
Universitas Multi Data Palembang**

Program Studi Informatika
Tugas Akhir Sarjana Komputer
Semester Genap Tahun 2022/2023

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP
INVESTASI KEUANGAN DI INDONESIA MENGGUNAKAN METODE
NAIVE BAYES**

Andreas Danny Agus Wahyudi 1822250060

Abstrak

Investasi keuangan merupakan hal yang penting bagi masyarakat Indonesia untuk mempersiapkan masa depan mereka secara finansial. Dalam era digital ini, media sosial seperti Twitter telah menjadi *platform* yang populer untuk berbagi opini dan pandangan tentang berbagai topik termasuk investasi keuangan. Dengan memanfaatkan data yang tersedia di *Twitter*, analisis sentimen dapat digunakan untuk memahami pandangan dan opini pengguna terkait investasi keuangan di Indonesia. Metode *Naive Bayes* dapat digunakan untuk melakukan analisis sentimen pada data Twitter dengan memanfaatkan teori probabilitas untuk mengklasifikasikan *tweet* yang berpandangan positif, negatif, atau netral tentang investasi keuangan di Indonesia. Jumlah data *tweet* tidak seimbang, sehingga perlu melakukan SMOTE *over-sampling* sehingga dataset tersebut seimbang dan melakukan pengujian menggunakan *k-fold validation* sehingga bisa melihat *confusion matrix* serta mendapatkan nilai akurasi, *precision*, *recall*, dan *f1-score*. Berdasarkan dari sentimen yang didapat dari media sosial Twitter menunjukkan bahwa pengguna media sosial Twitter memiliki sentimen positif terhadap investasi keuangan di Indonesia dengan total jumlah sentimen positif sebanyak 426 data dari total 1000 data *tweet*. Data yang tidak seimbang mempengaruhi hasil klasifikasi yaitu akurasi sebesar 45% dengan metode *up-sampling* SMOTE dan akurasi sebesar 89% tanpa menggunakan metode *up-sampling* SMOTE.

Kata kunci: Investasi, *Twitter*, *Naive Bayes*, SMOTE

BAB 1

PENDAHULUAN

Bab ini membahas mengenai latar belakang, rumusan masalah, ruang lingkup, tujuan dan manfaat, metodologi penelitian, dan sistematika penulis.

1.1 Latar Belakang

Investasi keuangan merupakan hal yang penting bagi masyarakat Indonesia untuk mempersiapkan masa depan mereka secara finansial. Namun, ada banyak faktor yang memengaruhi keputusan seseorang untuk berinvestasi, termasuk pandemi Covid-19 yang sedang terjadi, kondisi perekonomian global, dan kebijakan pemerintah terkait investasi.

Dalam era digital ini, media sosial seperti Twitter telah menjadi *platform* yang populer untuk berbagi opini dan pandangan tentang berbagai topik termasuk investasi keuangan. Dengan memanfaatkan data yang tersedia di Twitter, analisis sentimen dapat digunakan untuk memahami pandangan dan opini pengguna terkait investasi keuangan di Indonesia (SAPUTRA, 2020).

Investasi di Indonesia sendiri ada beberapa kejadian dengan investasi bodong yang sangat banyak terjadi. Investasi bodong adalah suatu bentuk investasi ilegal yang menjanjikan keuntungan besar dalam waktu singkat tanpa adanya risiko. Fenomena ini telah menjadi masalah serius di Indonesia dan banyak negara lainnya, karena telah menelan banyak korban dan merugikan banyak orang. Berikut contoh investasi bodong uang sering terjadi di Indonesia adalah janji keuntungan besar, Penawaran investasi dengan tahap awal rendah, promosi melalui media sosial, tidak memiliki ijin atau regulasi yang sah, dan sulitnya mendapatkan informasi yang akurat (SAPUTRA, 2020).

Ada beberapa algoritma yang dapat digunakan untuk analisis sentimen pada data Twitter, selain *Naive Bayes*. Beberapa di antaranya adalah: *Regresi Logistik (Logistic Regression)* adalah salah satu algoritma yang sering digunakan dalam analisis sentimen pada data Twitter. Algoritma ini bekerja dengan membuat model *regresi* yang dapat memprediksi probabilitas suatu tweet atau teks lainnya termasuk dalam kategori positif, negatif, atau netral. *Regresi*

Logistik memiliki keunggulan dalam menangani data yang bersifat kategorikal dan kontinu, serta mudah digunakan dan diinterpretasikan. (W et al., 2019).

Metode Mesin Pendukung Vektor (SVM) adalah algoritma yang dapat digunakan untuk mengklasifikasikan teks menjadi kelompok-kelompok yang diinginkan. SVM menggunakan teknik pembelajaran mesin untuk mengidentifikasi pola dalam data dan meningkatkan kinerja klasifikasi pada setiap iterasi. Dalam konteks analisis sentimen pada data Twitter, SVM dapat digunakan untuk mengklasifikasikan *tweet* menjadi kategori sentimen seperti positif, negatif, atau netral. (SAPUTRA, 2020).

Jaringan Saraf Tiruan (ANN) juga merupakan algoritma yang populer dalam analisis sentimen pada data Twitter. ANN dapat digunakan untuk mengklasifikasikan teks ke dalam kategori sentimen yang sama seperti SVM, yaitu positif, negatif, atau netral. Algoritma ini bekerja dengan memetakan input (teks) ke dalam output (kategori sentimen) menggunakan serangkaian lapisan *neuron* yang dipelajari dari data pelatihan. (SAPUTRA, 2020).

Salah satu metode yang umum digunakan dalam analisis sentimen pada data Twitter terkait investasi keuangan di Indonesia adalah metode *Naive Bayes*. Metode ini mengandalkan teori probabilitas untuk mengklasifikasikan teks ke dalam kategori positif, negatif, atau netral. Dalam analisis sentimen, metode *Naive Bayes* digunakan untuk mengklasifikasikan *tweet* yang mengungkapkan pandangan positif, negatif, atau netral tentang investasi keuangan di Indonesia. (Rafih Munir et al., 2022).

Perbandingan antara algoritma-algoritma tersebut dalam analisis sentimen pada data Twitter dapat dilakukan dengan membandingkan kinerja masing-masing algoritma dalam mengklasifikasikan data ke dalam kategori sentimen yang diinginkan. Kinerja algoritma dapat diukur dengan menggunakan beberapa metrik evaluasi, seperti akurasi, *presisi*, *recall*, dan *F1-score* (Amaliah & Dwi Nuryana, 2022).

Studi terbaru yang membandingkan kinerja beberapa algoritma dalam analisis sentimen pada data Twitter menunjukkan bahwa *Naive Bayes* memiliki performa yang cukup baik dibandingkan dengan algoritma lainnya, terutama ketika jumlah data pelatihan terbatas atau ketika data tidak seimbang (*imbalance class*). Namun, performa *Naive Bayes* juga tergantung pada kualitas data pelatihan dan pemilihan fitur yang tepat. Oleh karena itu, pemilihan algoritma terbaik untuk analisis sentimen pada data Twitter perlu dipertimbangkan secara cermat berdasarkan karakteristik data yang ada (Rafih Munir et al., 2022).

Hasil dari analisis sentimen ini dapat memberikan informasi yang berharga bagi investor dan pelaku pasar keuangan untuk memahami pandangan dan opini masyarakat terkait investasi keuangan di Indonesia. Dengan demikian, analisis sentimen pengguna Twitter terhadap investasi keuangan di Indonesia menggunakan metode *Naive Bayes* memiliki nilai penting dalam meningkatkan pemahaman tentang pasar keuangan Indonesia (Helmiah et al., 2020).

Pada penelitian sebelumnya yang dilakukan oleh Bimananda W, Insan Riski, Karina Dwi, Rani Nooraeni, Theresa Siahaan, dan Yosherina Dhea pada tahun 2019, mereka melakukan analisis *Text Mining* terhadap berita di Twitter yang berkaitan dengan infrastruktur di Indonesia. Metode klasifikasi yang digunakan dalam penelitian tersebut adalah *Naive Bayes* dengan tujuan untuk meningkatkan kesejahteraan masyarakat. Pada masa pemerintahan Presiden Joko Widodo, pembangunan infrastruktur dilakukan secara intensif guna mendukung pertumbuhan ekonomi dan peningkatan kesejahteraan masyarakat. Pembangunan infrastruktur ini memiliki dampak baik secara positif maupun negatif terhadap kehidupan masyarakat. Oleh karena itu, masyarakat sering memberikan pendapat dan komentar terkait pembangunan infrastruktur melalui berbagai *platform* media sosial, termasuk Twitter.

Hasil penelitian tersebut menunjukkan bahwa proporsi sentimen negatif lebih besar daripada sentimen positif dalam cuitan-cuitan terkait infrastruktur di Twitter. Melalui pengklasifikasian menggunakan metode *Naive Bayes*, didapatkan model terbaik pada kasus bandara dengan akurasi sebesar 82%, *presisi* sebesar 0,84, dan *recall* sebesar 0,48.

Penelitian yang dilakukan oleh Theda Yuwana Saputra pada tahun 2020 berjudul "Analisis Sentimen Berita Saham Menggunakan Metode *Naive Bayes*". Penelitian ini membahas tentang cara investasi dalam saham dan pengaruh sentimen berita terhadap pergerakan harga saham. Tujuan dari penelitian ini adalah membangun sistem yang memberikan informasi tentang sentimen berita saham untuk membantu investor dalam pengambilan keputusan investasi. Data yang digunakan dalam penelitian ini diambil dari situs berita saham kontan.com dan dilakukan preprocessing seperti case folding, stemming, penghapusan kata singkatan, kata penghubung, dan angka. Metode klasifikasi yang digunakan adalah *Naive Bayes* dengan fitur *TF-IDF* dan *TF-IDF Bigram*.

Hasil pengujian menunjukkan bahwa penggunaan metode *Naive Bayes* dalam klasifikasi berita saham memberikan akurasi rata-rata sebesar 76,79% untuk fitur *TF-IDF*, dan 77,13% untuk fitur *TF-IDF Bigram*. Pemilihan fitur sebesar 10% menghasilkan hasil *recall* yang terbaik dibandingkan dengan 25%, 50%, dan 100%. Semakin banyak fitur yang digunakan,

nilai *precision* meningkat tetapi nilai *recall* menurun. Hal ini disebabkan oleh *inklusi* fitur-fitur yang kurang penting.

Penelitian yang dilakukan oleh Muhammad Rafih S Munir, Hartatik, dan Andri Syafrianto pada tahun 2022 berjudul "Sentimen Analisis *Thread* Investasi Pada Twitter Menggunakan *Naïve Bayes*". Penelitian ini membahas tentang penggunaan media sosial sebagai sumber informasi tentang investasi dan bagaimana melakukan analisis sentimen terhadap *thread* investasi di Twitter. Tujuan dari penelitian ini adalah untuk mengidentifikasi *feedback* positif dan negatif terkait investasi menggunakan metode *Naïve Bayes*. Data yang digunakan berasal dari *thread-thread* investasi di Twitter, dan metode klasifikasi yang digunakan adalah *Naïve Bayes Multinomial*.

Hasil penelitian menunjukkan bahwa metode *Naïve Bayes Multinomial* dapat memberikan hasil dengan nilai *precision* sebesar 0.69, *recall* sebesar 0.70, *f1-score (f-measure)* sebesar 0.74, dan akurasi sebesar 0.7. Aplikasi yang dikembangkan dalam penelitian ini termasuk dalam bidang *opinion mining*, yaitu untuk mengetahui apakah suatu teks termasuk dalam kategori positif atau negatif. Proses penelitian dilakukan dengan mengambil *thread-thread* terkait investasi di Twitter, dan setiap kata dalam *thread* dianggap sebagai fitur penting dalam analisis sentimen. Metode *Naïve Bayes* merupakan metode *supervised learning* yang membutuhkan data *training*, dan dalam penelitian ini proses pelabelan data dilakukan secara manual oleh peneliti.

Oleh karena itu pada penelitian ini akan dilakukan analisis sentimen pengguna twitter tentang investasi keuangan di Indonesia dengan menggunakan metode *naive bayes*.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah dijelaskan, maka perumusan masalah dalam penelitian ini adalah bagaimana hasil analisis sentimen pengguna media sosial Twitter tentang investasi keuangan di Indonesia?

1.3 Ruang Lingkup

Ruang lingkup yang digunakan dalam penelitian analisis sentimen ini antara lain sebagai berikut:

1. Metode yang digunakan adalah metode *Naive Bayes*.
2. Data yang digunakan adalah data dari *tweet* pada media sosial Twitter.
3. Pengumpulan dataset twitter dilakukan dengan bantuan *library tweepy* sebagai penyedia layanan API twitter.

4. Data yang digunakan menggunakan hashtag (#) #investasi.
5. Data yang diperoleh adalah data tahun 2023.
6. Penelitian ini akan menggunakan teknik resampling data *SMOTE (Synthetic Minority Over-sampling Technique)* untuk menangani ketidakseimbangan kelas pada dataset yang akan digunakan.
7. Pengelompokan label data dibagi menjadi 3 kelompok yaitu data dengan label positif, label netral, dan data dengan label negatif.
8. Pemberian label dilakukan secara manual dengan bantuan dari ahli bahasa Indonesia yaitu Ibu Tresiana Sari Diah Utami, M.Pd dosen Universitas Katholik Musi Charitas.
9. *Library* yang digunakan sebagai alat bantu penelitian ini adalah NLTK dan Sastrawi.
10. Bahasa pemrograman yang digunakan sebagai alat bantu penelitian adalah *Python*.
11. *Tools* yang digunakan sebagai alat bantu penelitian adalah *Google Colab*.

1.4 Tujuan dan Manfaat

Tujuan dari penelitian ini adalah sebagai berikut :

1. Memahami cara menerapkan Metode *Naive Bayes* untuk menganalisis sentimen pada data Twitter.
2. Mendapatkan hasil analisa dan informasi yang dapat membantu pihak - pihak terkait dalam menyikapi keberadaan investasi keuangan di Indonesia saat ini.

Manfaat dari penelitian antara lain sebagai berikut :

1. Dengan metode *Naive Bayes* dapat diperoleh hasil analisa yang dapat meberikan informasi umum terkait investasi keuangan di Indonesia .
2. Bagi penulis, dapat memberikan pengetahuan bahwa apa yang terposting di sosial media dapat digunakan sebagai data analitik yang dapat menjadi informasi berharga.

1.5 Metode Penelitian

1. Identifikasi problem

Di tahapan ini dilakukan identifikasi dilema penelitian tentang analisis sentimen pengguna Twitter terhadap Investasi keuangan di Indonesia menggunakan metode *Naive Bayes*.

2. Studi Literatur

Di tahapan ini dilakukan pembelajaran literatur berupa jurnal dan buku yang berkaitan menggunakan topik penelitian ini agar mendapatkan berita dan data yang bisa digunakan buat penelitian.

3. Pengumpulan Dataset Twitter

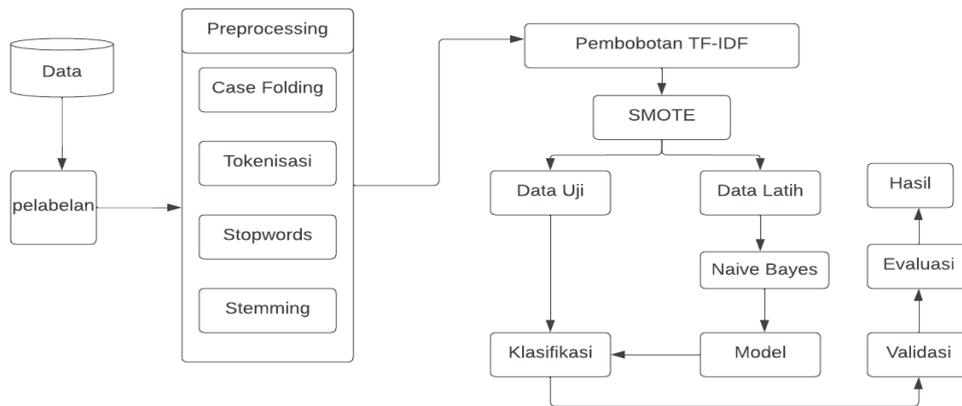
Di tahapan ini dilakukan pengumpulan data tweet menggunakan alat bantu yaitu dengan *Google Colab*, bahasa pemrograman yang dilakukan merupakan *Python*, serta *Library Tweepy*. buat bisa memakai *Libray Tweepy* tersebut diharuskan buat mempunyai akun Twitter dahulu serta sudah mendaftarkan akun tersebut sebagai akun *developer* buat menerima *access key* yang dibutuhkan buat melakukan pengumpulan data Twitter.

	datetime	username	content
0	2023-04-16 11:24:19+00:00	ilovember	@worksfess 1. Bisa. Tp harus belajar dulu mengenai investasi. Tentukan tujuan keuangan dan jgn inves berdasarkan apa kata orang tp harus keputusan sendiri2. Udah luar biasa nder menurutku utk bisa nabung segitu padahal dg penghasilan magang. Kamu keren sobat3. Menurutku sih kamu fokus ke—
1	2023-04-16 11:10:50+00:00	dawiguna	Siapapun yg mengambil produk keuangan, entah investasi, entah asuransi/n/nanti WAKTU MEMBUKTIKAN kebenaran analisa dan solusi/n/nDengan persetujuan Nasabah di depan, maka semua akibat menjadi 100% tanggungan Nasabah/n/nWalaupun ada bujuk rayu dari agen, tetap Nasabah sudah setuju ya
2	2023-04-16 11:10:47+00:00	dawiguna	Dunia kerja di Indonesia harus ditata/n/nMisalnya, banyak agen asuransi, agen bursa berjangka, bahkan agen bank, n/nyg ditugaskan mendapatkan Nasabah keuangan, namun mereka tidak sepenuhnya memahami produk yg mereka jual/n/nTidak paham ekonomi, iklim investasi, tingkat risiko, dsb
3	2023-04-16 05:10:37+00:00	bentangpustaka	Il Investor wajib tahu l/n/nKamu udah wajib menunaikan zakat saham nggak nih? n/nyuk, ngobrol bareng Mintang di kolom komentar/n/nBuat kamu para investor pemula, yuk, pelajari seluk beluk dunia keuangan dan investasi/n/nmelalui buku A Quick Guide for Investors! https://t.co/WbvVFIEuX7n
4	2023-04-16 01:16:45+00:00	frrick	@itznyaaw @2byfess setelahnya baru investasi berdasarkan tujuan keuangan yg mau dicapai, buat instrumen investasinya yg cocok buat pemula dan modal terjangkau ambil reksadana, nanti pas awal ambil reksadana pasar uang dulu aja sambil belajar, aku ada threadnya kalau mau baca/n/n https://t.co/XTHd2oFmsQ
...
995	2023-02-13 11:56:47+00:00	NekoSamaDesu96	@GemaShowindo Kurangnya literasi tentang keuangan dan how to invest adalah salah satu penyebab kejadian seperti ini. Salah satunya adalah keluarganya teman ane, padahal anaknya sudah melek tentang keuangan dan investasi, tapi ketika ngasih tau ortunya tentang ini, dibilang 'sok tau'...
996	2023-02-13 11:09:30+00:00	1001ramelan	Anda mungkin akan harus berurusan dengan kerabat yang lebih tua hari ini. Luangkan waktu memeriksa kesepakatan tempat tinggal atau investasi jangka panjang lainnya. Peluang di bidang keuangan akan berkembang.
997	2023-02-13 11:01:36+00:00	imputoun	Jaga hati, jaga pandangan :) nana tuh udah jdi best standard asiiii. Knowledge dia tentang investasi, keuangan, urusan rumah dll udah gausah ditanya. Tapi sayang... semua ini terhalang babi goreng :(
998	2023-02-13 09:02:45+00:00	CommBank_ID	#investisfun #investasi #investment #yukinvestasi #saham #reksadana #KendalikanUangmu #keuangan #SmartWealth #money #financial #financialfreedom #financialplanning #growth #lot #melekinvestasi #reksadana #produk #saham #ojk

Gambar 1.1 Dataset pengambilan data asal Twitter

4. Perancangan Sistem

Tahapan ini dilakukan perancangan sistem mulai asal pembobotan *TF-IDF*, dimana dilakukan proses perubahan kata sebagai bentuk nomor atau vector, sedangkan *TF-IDF* digunakan buat menentukan frekuensi sebuah istilah didalam sebuah dokumen serta anugerah bobot pada setiap istilah di setiap dokumen buat mencari serta menghitung beberapa kali istilah itu timbul. Penelitian ini memakai metode *Naive Bayes*, diawali dengan diinput data yang telah dibagi menjadi 80% data latih dan 20% data uji.



Gambar 1.2 Bagan Perancang Sistem

5. Validasi

pada tahapan ini dilakukan validasi contoh, Validasi yang dimaksud adalah tahapan buat pengecekan apakah akibat yang di bisa dari sistem, dapat di pastikan benar serta telah sinkron dengan yang diinginkan. di termin validasi data latih ini memakai metode *K-Fold (10 fold) Cross Validation*.

6. Evaluasi

tahap ini melakukan penilaian terhadap sistem yang sudah dirancang. Tujuan asal tahapan ini merupakan memastikan kebenaran akibat yang didapat serta meminimalisir kesalahan. di termin evaluasi ini juga digunakan metode *Confusion matrix* yang akan menghitung *precision*, *recall*, *F1-score* serta *accuracy*.

1.6 Sistematika Penulis

Sistem penulisan laporan tugas akhir yang akan dilakukan sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini membahas mengenai latar belakang, rumusan masalah, ruang lingkup, tujuan dan manfaat, metodologi penelitian, dan sistematika penulis.

BAB 2 LANDASAN TEORI

Bab ini menjelaskan tentang teori-teori dasar yang berkaitan dengan penelitian analisis sentimen, metode Algoritma *Naive bayes*, dan penelitian terdahulu.

BAB 3 METODE PENELITIAN

Bab ini menjelaskan tentang kebutuhan perangkat keras, perangkat lunak, serta tahapan dari proses penelitian yang akan dilakukan.

BAB 4 HASIL DAN PEMBAHASAN

Bab ini membahas mengenai dari tahapan yang akan dilakukan dan menjelaskan hasil implementasi dari metode yang dipakai yaitu metode Algoritma *Naive Bayes* serta dari pengujian.

BAB 5 PENUTUP

Bab ini menjelaskan tentang kesimpulan yang telah diperoleh dari penelitian yang telah dilakukan dan saran-saran yang dapat bermanfaat bagi para penelitian lain sebagai penelitian selanjutnya.

DAFTAR PUSTAKA

- Amaliah, F., & Dwi Nuryana, I. K. (2022). Perbandingan akurasi metode lexicon based dan naive bayes classifier pada analisis sentimen pendapat masyarakat terhadap aplikasi investasi pada media twitter. *Journal of Informatics and Computer Science (JINACS)*, 3(03), 384–393. <https://doi.org/10.26740/jinacs.v3n03.p384-393>
- Anggraeni, D. T. (2022). Analisis sentimen vaksinasi booster covid-19 pada platform twitter menggunakan metode naïve bayes. *EXPERT: Jurnal manajemen sistem informasi dan teknologi*, 12(2), 113. <https://doi.org/10.36448/expert.v12i2.2812>
- Azhar, R., Surahman, A., & Juliane, C. (2022). Analisis sentimen terhadap cryptocurrency berbasis python textblob menggunakan algoritma naïve bayes. *Jurnal sains komputer & informatika (J-SAKTI)*, 6(1), 267–281.
- Guntur, M., Santony, J., & Yuhandri, Y. (2018). Prediksi harga emas dengan menggunakan metode naïve bayes dalam investasi untuk meminimalisasi resiko. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 2(1), 354–360. <https://doi.org/10.29207/resti.v2i1.276>
- Haryadi, D., & Mandala, R. (2019). Prediksi harga minyak kelapa sawit dalam investasi dengan membandingkan algoritma naïve bayes, support vector machine dan k-nearest neighbor. *IT for society*, 4(1). <https://doi.org/10.33021/itfs.v4i1.1181>
- Helmiah, N., Nooraeni, R., Nulkarim, A. R., Munia, N., Susanti, A., Putra, A. P., & Taufiqurrahman, F. (2020). Penerapan metode naïve bayes dalam analisis persepsi masyarakat mengenai rencana pengesahan RUU omnibus law di bidang investasi dan ketenagakerjaan tahun 2020 di Indonesia. *Jurnal MSA (matematika dan statistika serta aplikasinya)*, 8(2), 48. <https://doi.org/10.24252/msa.v8i2.16743>
- Leonie Syafira, & Rikumahu, B. (2020). Analisis korelasi sentimen pada twitter terhadap abnormal return saham (studi kasus pada saham indeks Lq45 di twitter). *Jurnal mitra manajemen*, 4(9), 1322–1335. <https://doi.org/10.52160/ejmm.v4i9.458>
- Niyah, T. Y., & Yamasari, Y. (2021). NBC berbasis ekstraksi fitur untuk analisis sentimen masyarakat terhadap fluktuasi harga saham. *Journal of Informatics and Computer Science (JINACS)*, 3(02), 113–122. <https://doi.org/10.26740/jinacs.v3n02.p113-122>
- Pribadi, M. R., Purnomo, H. D., Hendry, Hartomo, K. D., Sembiring, I., & Iriani, A. (2022). Improving the accuracy of text classification using the over sampling technique in the case of sinovac vaccine. *International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2022-October(October), 106–110. <https://doi.org/10.23919/EECSI56542.2022.9946508>
- Rafih Munir, M. S., Syafrianto, A., & Ilmu Komputer, F. (2022). Sentimen analis thread investasi pada twitter menggunakan naïve bayes. *IJCSR: The Indonesian Journal of*

Computer Science Research, 1, 44–52. <https://subset.id/index.php/IJCSR>

Rusardi, M. N., Rahayudi, B., & Adikara, P. P. (2022). Analisis sentimen masyarakat terhadap isu new normal scenario berdasarkan opini dari twitter menggunakan algoritma naive bayes classifier. *Jurnal pengembangan Teknologi Informasi Dan Ilmu Komputer*, 6(3), 1434–1440.

SAPUTRA, T. (2020). Analisis sentimen berita saham menggunakan algoritma naive bayes. *Katalog.Ukdw.Ac.Id*. <http://katalog.ukdw.ac.id/id/eprint/4166>

W, B., Riski, I., Dwi, K., Nooraeni, R., Siahaan, T., & Dhea, Y. (2019). Analisis text mining dari cuitan twitter mengenai infrastruktur di Indonesia dengan metode klasifikasi naïve bayes. *Eigen mathematics journal*, 2(2), 92–101. <https://doi.org/10.29303/emj.v1i2.36>

