PRISMA 8 (2025): 407-413
PRISMA, Prosiding Seminar Nasional Matematika
https://journal.unnes.ac.id/sju/index.php/prisma/
ISSN 2613-9189

Dinamika Sentimen Publik dalam Suksesi Pemerintahan Indonesia berdasarkan Analisis Data Media Sosial

Galih Kusuma Wijaya[a,*], Adelia Venie Diniar[b], Shata Alwan Jalaluddin[c], Iqbal Kharisudin[d]

[a,b,c,d] *Statistika dan Sains Data FMIPA UNNES, Semarang, Indonesia*

** Alamat Surel: galihkjaya15@students.unnes.ac.id*

Abstrak

Dinamika sentimen publik terhadap suksesi pemerintahan di Indonesia menjadi aspek penting dalam memahami persepsi masyarakat yang tercermin melalui sosial media. Penelitian ini bertujuan untuk mengembangkan model klasifikasi sentimen menggunakan model *RandomForest* untuk menganalisis *tweet* terkait suksesi pemerintahan Indonesia. Data yang digunakan terdiri dari 637 *tweets* yang sudah melalui proses *Data Cleaning*. Model dilatih menggunakan dataset yang dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Berdasarkan hasil pengujian, model memperoleh akurasi sebesar 66%, menunjukkan bahwa model cukup mampu dalam mengklasifikasikan sebagian besar data secara akurat. Analisis lebih lanjut menunjukkan bahwa model memiliki performa terbaik dalam mendeteksi sentimen Netral, dengan nilai *precision* sebesar 0,63, *recall* sebesar 0,86, dan *F1-score* sebesar 0,73. Namun, performa untuk kelas Negatif dan Positif masih tergolong kurang optimal, terlihat dari nilai *recall* yang rendah, masing-masing sebesar 0,45 untuk Negatif dan 0,56 untuk Positif. Hal ini mengindikasikan bahwa model masih sering salah dalam mengklasifikasikan data Negatif dan Positif, yang kemungkinan disebabkan oleh ketidakseimbangan data yang ada. Penelitian ini diharapkan dapat memberikan informasi mengenai bagaimana sentimen publik terhadap suksesi pemerintahan di Indonesia saat ini.

Kata kunci: analisis sentimen, suksesi pemerintahan Indonesia, crawling data, text mining

© 2025 Dipublikasikan oleh Jurusan Matematika, Universitas Negeri Semarang

1. Pendahuluan

Tahun 2024 menjadi tahun suksesi pemerintahan indonesia dari pemerintahan Presiden Joko Widodo ke Presiden selanjutnya yaitu Presiden Prabowo. Suksesi adalah proses pergantian kepemimpinan sesuai dengan peraturan perundang-undangan yang berlaku (KBBI, 2023). Suksesi menjadi salah satu momen krusial dalam perjalanan politik suatu negara, termasuk Indonesia. Suksesi bukan sekedar mengganti kepala pemerintahan. Menurut Surbekti (1998), mengganti seorang kepala pemerintahan juga berarti mengganti pihak yang ikut memerintah, seperti para menteri, parapenjabat pemerintah yang diangkat oleh kepala pemerintahan, kepala berbagai instansi, seluruh jajaran birokrasi, dan kelompok-kelompok masyarakat yang diuntungkan secara ekonomi, keagamaan, ideologik dan kultural oleh kekuasaan tersebut. Sehingga proses ini akan mengubah dinamika politik, sosial, dan ekonomi.

Perubahan dinamika dan kebijakan setelah pergantian kepemimpinan menjadi sorotan utama bagi masyarakat. Dampak dari perubahan tersebut dirasakan oleh banyak orang, sehingga mereka aktif menyuarakan pendapat mereka. Ada yang memberikan pujian atas kebijakan baru, sementara yang lain menyampaikan keluhan dan kritik terhadap kebijakan yang dianggap kurang memuaskan. Masyarakat berharap suara mereka didengar dan dipertimbangkan dalam proses pengambilan keputusan oleh pemerintahan yang baru.

Dalam era digital saat ini, internet telah menjadi bagian integral dari kehidupan sehari - hari kita, dan jutaan orang di seluruh dunia mengaksesnya secara aktif (Keen, 2015). *Platform* seperti Facebook, Twitter, Instagram, dan LinkedIn telah mengubah cara kita berinteraksi dengan teman, keluarga,

To cite this article:
Wijaya, Diniar, Jalaluddin, & Kharisudin (2025). Dinamika Sentimen Publik dalam Suksesi Pemerintahan Indonesia berdasarkan Analisis Data Media Sosial. *PRISMA, Prosiding Seminar Nasional Matematika 8*, 407-413

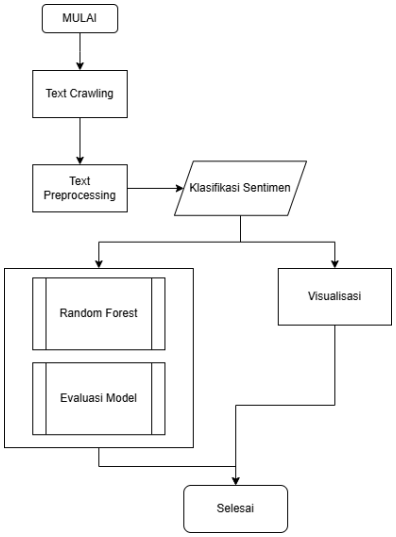
dan kolega (Lipschultz, 2020). *Platform* tersebut juga digunakan masyarakat untuk mengekspresikan pandangan dan sentimennya, salah satu platform yang sering digunakan adalah X (Twitter). Tanggapan tersebut diklasifikasikan menjadi positif, negatif, dan netral menggunakan analisis sentimen. Analisis sentimen adalah studi komputasi opini, sikap, dan emosi orang terhadap suatu entitas (Medhat et al., 2014). Analisis sentimen, yang juga dikenal sebagai penambangan opini juga diartikan sebagai salah satu aplikasi NLP (*Natural Language Processing*) yang paling banyak digunakan untuk mengidentifikasi emosi manusia dari ulasan mereka (Shaik et al., 2022). Memahami tren dan dinamika komunikasi digital merupakan hal yang penting agar dapat mengatasi tantangan dan memanfaatkan peluang yang ada (Andzani & Irwansyah, 2023).

Memahami sentimen publik dapat membantu pemerintahan untuk merumuskan kebijakan yang baik untuk dilakukan kedepannya sesuai dengan pendapat masyarakat yang telah dianalisis. Penelitian ini akan menganalisis sentimen publik dalam Suksesi Pemerintahan. Metode yang digunakan untuk menganalisis opini publik dengan data yang diperoleh pada sosial media X tersebut dilakukan dengan mengembangkan model klasifikasi sentimen menggunakan model *Machine Learning* dengan pendekatan *RandomForest*. Random Forest merupakan algoritma dalam klasifikasi data supervised, menggabungkan beberapa *tree*, yang masing - masing dilatih secara terpisah dimana model dasar dilatih dan dikombinasikan menggunakan skema pembobotan yang canggih, biasanya *tree* dilatih secara independen dan prediksi tree digabungkan melalui rata - rata (Denil et al., 2014).

2. Metode

Penelitian ini berbasis data kualitatif, data yang kami gunakan berasal dari media sosial X yang diambil dengan menggunakan metode *Text Crawling* dan dibantu dengan alat *Tweet Harvest.*

Dari proses *text crawling* ini, kami berhasil mendapatkan 637 *tweets* dari X dengan keyword "Suksesi Pemerintahan". Adapun tahapan penelitian yang akan kami lakukan dapat dilihat pada gambar berikut:



Gambar 1. Tahapan penelitian.

2.1. Text Mining

Text Mining adalah proses untuk memperoleh informasi dari sebuah teks (Deolika, Kusrini, et al., 2019). Proses pengumpulan data berbasis *Text Mining* kami lakukan dengan menggunakan alat *Tweet Harvest* (helmisatria, n.d.) yang bisa didapatkan melalui *https://github.com/helmisatria/tweet-harvest*, Kami menggunakan *keyword* "Suksesi Pemerintahan" dan menginput "id" postingan terkait untuk diambil data yang ada pada postingan tersebut, seperti komentar. Data yang sudah didapatkan kemudian akan tersimpan dalam bentuk *dataframe*.

2.2 Text Preprocessing & Klasifikasi Sentimen

Data yang berhasil didapatkan kemudian dilakukan proses *Text Cleaning* yang merupakan cara untuk menghilangkan karakter khusus (simbol), *retweet* (RT), *username*, dan *link* yang ada pada dataset (Effendi et al., 2024). Proses *Text Cleaning* pada penelitian ini memanfaatkan kamus kata baku yang diperoleh dari GitHub. Tahapan ini sangatlah krusial untuk menghindari kesalahan prediksi pada model akibat kata-kata yang salah tulis atau *typo.*

Pada tahapan tersebut, teks diubah menjadi *lowercase* untuk memastikan bahwa setiap kata direpresentasikan dalam satu bentuk, sehingga membantu model untuk mengklasifikasikan data secara lebih akurat. Seperti contohnya kata “Bagus” dan “bagus”, mungkin secara semantik memiliki arti yang sama, namun dalam algoritma tidak demikian. Kata “Bagus” biasanya akan diklasifikasikan menjadi sentimen netral karena memiliki arti nama seseorang.

Setelah melalui tahap *Text Cleaning*, langkah selanjutnya adalah tokenisasi. Tokenisasi adalah proses pemotongan kalimat menjadi kata per kata (Surbakti et al., 2021). Proses ini bertujuan untuk mempermudah model dalam mengklasifikasikan sentimen pada sebuah kalimat. Dalam tokenisasi, setiap kata ataupun frasa diidentifikasi sebagai fitur unik yang kemudian dapat dihitung frekuensi kemunculan kata tersebut dan dipelajari polanya oleh model.

Tabel 1. *Hasil text cleaning dan tokenisasi.*

Tweet	Tweet_koreksi
[Jokowi, presiden, ngibul]	[[Jokowi], [presiden], [berbohong]]
[sorry, ye]	[[maaf], [ye]]

Data yang sudah bersih dan ditokenisasi kemudian dapat digunakan untuk klasifikasi sentimen. Penelitian ini menggunakan kamus sentimen untuk membantu model mengklasifikasikan kata ke dalam sentimen tertentu. Ada tiga kelas yang diklasifikasikan, yaitu: netral, positif, dan negatif. Contoh kata yang berada dalam masing-masing kamus sentimen adalah seperti berikut:

Tabel 2. *Kamus sentimen.*

Negatif	Positif
leceh	akrab
hina	aktif
kasar	aktualisasi
kalah	akur
nista	akurat

Kata yang tidak terdapat dalam kedua kamus tersebut kemudian akan diklasifikasikan ke dalam kelas netral. Berikut adalah contoh hasil klasifikasi sentimen:

Tabel 3. *Hasil klasifikasi sentimen.*

Kata	Sentimen
terimakasih jokowi	Positif
presiden	Netral
serakah	Negatif
usak	Netral
demokrasi	Positif

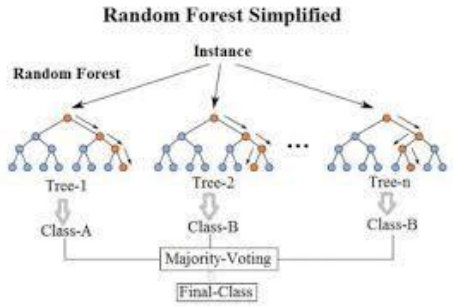
Pada **Tabel 3**, terlihat sentimen dari tiap kata perkata. Namun belum tentu kalimatnya akan memiliki sentimen yang sama. Dalam klasifikasi sentimen pada kalimat, model akan memperhatikan pola yang ada pada tiap kata, seperti kata “sangat” yang memiliki sentimen netral, jika dipasangkan dengan kata “buruk” maka akan membentuk frasa “sangat buruk” yang memiliki sentimen negatif.

2.3. Random Forest

Penelitian ini menggunakan metode *Random Forest. Random Forest* adalah model klasifikasi yang terdiri dari sejumlah *decision tree* dan tiap *decision tree* menggunakan subset *bootstrap* yang diambil dari sampel melalui pemilihan acak. (Heryadi & Wahyono, 2020). Pengambilan keputusan pada *RandomForest* ditetapkan dengan persamaan:

$$\boldsymbol{y(X) = mode\{C_1(x), C_2(x), \dots, C_m(x)}$$

dimana **m** mewakili jumlah pohon atau *decision tree* dalam model *RandomForest*, dan **mode** adalah operator untuk memilih input bilangan dengan frekuensi tertinggi (Heryadi & Wahyono, 2020). Respon suatu *decision tree* diprediksi dengan menggabungkan hasil prediksi *m* pohon, kemudian hasil dari prediksi terbanyak akan menjadi *output* model ini (Dewi et al., 2011).



Gambar 2. *Visualisasi RandomForest.*

Proses pelatihan dengan menggunakan *Random Forest* diawali dengan membagi dataset menjadi 8:2, dimana 80 persen untuk pelatihan dan 20 persen untuk pengujian. Pembagian dataset menjadi data pelatihan dan pengujian pada model *Random Forest* adalah untuk membentuk regresi yang nantinya akan digunakan model untuk melakukan prediksi (Susetianingtias et al., 2022) Selanjutnya kedua data akan divektorisasi dengan menggunakan *TF-IDF* untuk menghitung frekuensi relatif dari suatu kata dalam sebuah data atau kelompok data (Mentari, 2023).

Proses pelatihan ini dilakukan dalam beberapa *epoch. Epoch* sendiri merupakan proses iterative di mana model belajar dari data dengan mengulang pelatihan pada seluruh dataset. Semakin banyak *epoch*, maka model akan semakin memahami pola pada data dan menghasilkan akurasi yang tinggi (Wasil, 2022). Namun, hal ini perlu diperhatikan karena jika *epoch* yang diberikan terlalu banyak, maka model kemungkinan akan mengalami *overfitting*, di mana model terbiasa dengan data pelatihan dan sulit beradaptasi dengan data baru.

Tabel 4. *Perbandingan akurasi sesuai epoch.*

Model	Akurasi
100 Epoch	0.65625
200 Epoch	0.69531

3. Hasil dan Pembahasan

Hasil dari analisis sebelumnya memberikan informasi yang lebih mendalam terkait sentimen masyarakat terhadap topik yang dibahas. Dengan mengelompokkan sentimen masyarakat menjadi sentimen positif, negatif, dan netral, kita bisa memahami pola respons publik secara umum dan mengidentifikasi persepsi mayoritas terhadap suksesi pemerintahan ini.

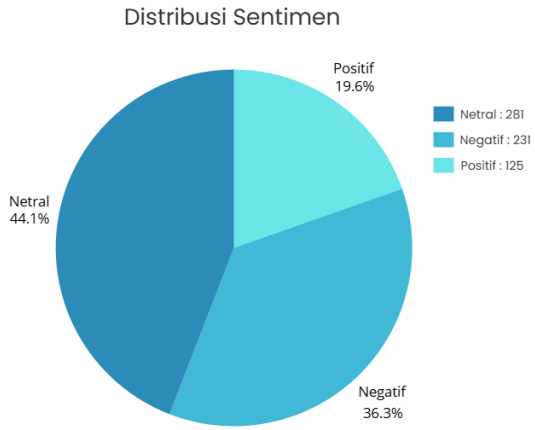
(1) *Sentimen Positif*, menunjukkan aspek-aspek dalam pemerintahan sebelumnya yang diapresiasi masyarakat, dan menunjukkan harapan yang tinggi terhadap pemerintahan yang baru.

(2) *Sentimen Negatif*, berperan sebagai sumber masukkan bagi pemerintahan untuk bisa mengevaluasi kinerja pada masa pemerintahan sebelumnya, dan bisa menjadi fokus perbaikan bagi pemerintahan yang baru.

(3) Sentimen Netral, biasanya berisikan mengenai opini-opini masyarakat yang tidak bersifat emosional atau pernyataan informatif.

3.1. Analisis Deskriptif

Dalam pembahasan ini, penelitian ini menggunakan metode statistika deskriptif yang merupakan bagian dari statistika yang mempelajari cara pengumpulan dan penyajian data sehingga mudah dipahami (Irhamah et al., 2020). Untuk merepresentasikan sentimen dari data, *pie chart* dipilih karena merupakan salah satu grafik yang paling efisien dan dapat dengan mudah dipahami.



Gambar 3. *Pie chart distribusi sentimen.*

Pada **Gambar 3** bisa kita perhatikan bahwa mayoritas sentimen masyarakat adalah Netral. Namun yang perlu disorot adalah distribusi dari sentimen Positif yang hanya 19.6 persen dibandingkan sentimen Negatif yang memiliki distirbusi 36.3 persen. Hal ini menunjukkan bahwa masih banyak masyarakat yang menganggap suksesi pemerintahan kali ini gagal.

Gambar 4. *Wordcloud.*

Visualisasi selanjutnya digambarkan dengan menggunakan *wordcloud.* Dengan menggunakan *wordcloud,* gambaran frekuensi kata-kata dapat ditampilkan dalam bentuk yang menarik namun tetap informatif. Ukuran gambar teks dalam *wordcloud* menyesuaikan dengan frekuensi data, semakin banyak frekuensi kata digunakan, semakin besar pula ukuran kata tersebut ditampilkan di *wordcloud* (Pradana, 2020). Dalam **Gambar 4** kata yang memiliki frekuensi kemunculan tertinggi didominasi oleh kata yang memiliki sentimen negatif, seperti “jilat”, “buzzer”, “bodoh”, dan lainnya.

3.2. Random Forest

Hasil penelitian menggunakan *Random Forest* untuk analisis lanjut. Model ini digunakan untuk menyelesaikan masalah yang melibatkan regresi dan klasifikasi (Jlifi et al., 2024) Itulah mengapa *Random Forest* sangat cocok untuk analisis sentimen. Model ini membangun banyak pohon keputusan dan kemudian menggabungkannya untuk memberikan hasil yang akurat dan stabil, yang umumnya digunakan untuk klasifikasi (Zhang et al., 2022).

- Akurasi **66%**

Tabel 5. *Hasil uji RandomForest.*

	Precision	Recall	f1-score	Support
Negatif	0.73	0.45	0.56	42
Netral	0.63	0.86	0.73	59
Positif	0.71	0.56	0.62	27

Hasil uji *RandomForest* pada penelitian ini menunjukkan bagaimana model mengklasifikasikan sentimen dalam tiga kategori: Negatif, Netral, dan Positif. Seperti yang terlihat dalam **Tabel 5** hasil uji *RandomForest* mendapatkan skor akurasi 66% yang berarti model masih tergolong mampu untuk mengklasifikasikan sebagian data yang ada. Akurasi mengukur sejauh mana model dapat mengklasifikasikan data dengan benar secara keseluruhan, sedangkan *precision* mengukur sejauh mana prediksi tiap kelas sentimen yang benar. *Recall* mengukur sejauh mana model dapat mendeteksi dengan benar kelas sentimen, sedangkan *F1-score* merupakan perpaduan antara *precision* dan *recall* (Fadli & Saputra, 2023).

Pada kelas Negatif, model tidak cukup andal dalam mendeteksi kata yang memiliki sentimen Negatif, terlihat dari rendahnya nilai *recall.* Ini berarti ada sejumlah besar kata bersentimen Negatif yang diprediksi sebagai kelas lain, yang dapat disebabkan oleh ketidakseimbangan data atau fitur yang kurang mendukung unutk kelas sentimen Negatif ini. Sedangkan pada kelas Netral, model memiliki performa terbaik dalam mendeteksi kelas Netral, dengan *recall* tinggi yang menunjukkan bahwa model jarang salah dalam mengklasifikasikan data Netral.

Kemudian untuk kelas Positif, model menunjukkan kinerja yang lebih baik jika dibandingkan dengan kelas Negatif, tetapi masih bisa ditingkatkan. *Precision* yang cukup tinggi menunjukkan bahwa prediksi untuk kelas Positif cukup dapat diandalkan, namun *recall* yang lebih rendah mengindikasikan ada beberapa data Positif yang mungkin terklasifikasi ke kelas lain.

4. Simpulan

Hasil penelitian ini menyimpulkan bahwa sentimen publik terhadap suksesi pemerintahan Indonesia di tahun 2024 mayoritas bersifat netral dengan nilai sebesar 44.1% dari keseluruhan data. Penelitian ini juga mengidentifikasi banyaknya sentimen negatif yang lebih besar daripada sentimen positif dengan nilai 36.3% dibandingkan sentimen positif yang hanya sebesar 19.6%, hal ini membuktikan banyaknya masyarakat yang tidak puas terhadap proses transisi kepemimpinan dari Presiden saat ini ke Presiden terpilih. Dominasi kelas sentimen netral menunjukkan bahwa banyak masyarakat yang berhati-hati dalam menyuarakan pendapatnya terhadap topik-topik yang sensitif seperti suksesi pemerintahan ini.

Daftar Pustaka

Andzani, D., & Irwansyah. (2023). Dinamika Komunikasi Digital: Tren, Tantangan, dan Prospek Masa Depan. Jurnal Syntax Admiration, 4, 1964–1976. https://doi.org/10.46799/jsa.v4i11.743

Deolika, A., Kusrini, & Luthfi, E. T. (2019). ANALISIS PEMBOBOTAN KATA PADA KLASIFIKASI TEXT MINING. Jurnal Teknologi Informasi, 3(2).

Denil, M., Matheson, D., & De Freitas, N. (2014). Narrowing the Gap: Random Forests In Theory and In Practice. Proceedings of the 31st International Conference on Machine Learning, 32(1), 665–673.

Dewi, N. K., Syafitri, U. D., & Mulyadi, S. Y. (2011). PENERAPAN METODE RANDOM FOREST DALAM DRIVER ANALYSIS (The Application of Random Forest in Driver Analysis) (Vol. 16, Issue 1).

Effendi, M. M., Zy, A. T., & Arwan, A. (2024). Analisis Sentimen Media Sosial Twitter Terhadap Calon Presiden RI Tahun 2024 Menggunakan Klasifikasi Algoritma Naïve Bayes. https://doi.org/10.47065/josyc.v5i3.5210

Fadli, M., & Saputra, R. A. (2023). KLASIFIKASI DAN EVALUASI PERFORMA MODEL RANDOM FOREST UNTUK PREDIKSI STROKE Classification And Evaluation Of Performance Models Random Forest For Stroke Prediction. Jurnal Teknik, 12. http://jurnal.umt.ac.id/index.php/jt/index

Heryadi, Y., & Wahyono, T. (2020). Machine Learning: Konsep dan Implementasi. PENERBIT GAVA MEDIA.

Irhamah, Rakhmawati, N. A., & Nurhadi, H. (2020). Pengembangan Sistem Informasi Sederhana untuk Pengelolaan dan Pengolahan Data Tol Laut PT. PELNI (Persero) Cabang Surabaya. In Jurnal Pengabdian kepada Masyarakat-DRPM ITS (Vol. 4, Issue 2). https://nasional.kontan.co.id/,

Jlifi, B., Abidi, C., & Duvallet, C. (2024). Beyond the use of a novel Ensemble based Random Forest-BERT Model (Ens-RF-BERT) for the Sentiment Analysis of the hashtag COVID19 tweets. Social Network Analysis and Mining. https://doi.org/10.1007/s13278-024-01240-x

Keen, A. (2015). The Internet Is Not the Answer. Atlantic Monthly Press.

Lipschultz, J. H. (2020). Social Media Communication: Concepts, Practices, Data, Law and Ethics (2nd ed.). Routledge.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093–1113.

Mentari, Y. N. (2023). Analysis of the Impact of Vectorization Methods on Machine Learning-Based Sentiment Analysis of Tweets Regarding Readiness for Offline Learning (Vol. 11, Issue 2).

Pradana, M. G. (2020). PENGGUNAAN FITUR WORDCLOUD DAN DOCUMENT TERM MATRIX DALAM TEXT MINING. https://ejournal.upbatam.ac.id/index.php/jif

Shaik, S. A., Soni, M., & Singh, A. (2022). Sentiment Analysis: An Overview of Methods and Applications. International Journal of Data Science and Analysis, 8(2), 45–53

Surbakti, A. Q., Hayami, R., & Al-Amien, J. (2021). Analisa Tanggapan Terhadap Psbb Di Indonesia Dengan Algoritma Decision Tree Pada Twitter. Jurnal CoSciTech (Computer Science and Information Technology), 2(2), 91–97. https://doi.org/10.37859/coscitech.v2i2.2851

Surbakti, R. (1998). Reformasi lembaga kepresidenan RI: Suatu pemikiran tentang mekanisme suksesi dan pembatasan kekuasaan presiden [Pidato pengukuhan Guru Besar]. Universitas Airlangga.

Susetianingtias, D. T., Patriya, E., & Rodiah. (2022). Model Random Forest Regression Untuk Peramalan Penyebaran Covid-19 Di Indonesia. Decode: Jurnal Pendidikan Teknologi Informasi, 2(2), 84–95. https://doi.org/10.51454/decode.v2i2.48

Wasil, M., Harianto, & Fathurrahman. (2022). Pengaruh Epoch pada Akurasi menggunakan Convolutional Neural Network untuk Klasifikasi fashion dan Furniture. Jurnal Informatika Dan Teknologi, 5(1), 53. https://doi.org/10.29408/jit.v5i1.4393

Zhang, C., Ren, J., Liu, F., Li, X., & Liu, S. (2022). Three-way selection random forest algorithm based on decision boundary entropy. Applied Intelligence. https://doi.org/10.1007/s10489-021-03033-7