

Analisis Sentimen Ciutan *Twitter* Terkait Penerapan Permendikbudristek Nomor 30 Tahun 2021 Menggunakan *TextBlob* dan *Support Vector Machine*

Yesy Afrillia^{1✉}, Lidya Rosnita², Deassy Siska³

^{1,2} Prodi Teknik Informatika, Jurusan Informatika, Fakultas Teknik, Universitas Malikussaleh, Indonesia

³ Prodi Pendidikan Fisika, Jurusan Pendidikan Ilmu Alam, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Malikussaleh, Indonesia

Informasi Artikel

Riwayat Artikel

Diserahkan : 11-10-2022

Direvisi : 16-10-2022

Diterima : 20-10-2022

Kata Kunci:

Gender, Twitter, Permendikbudristek, PPKS.

Keywords :

Gender, Twitter, Permendikbudristek, PPKS

ABSTRAK

Di era yang serba digital segala aspek penerapan suatu kebijakan oleh pemerintah menjadi topik yang banyak dibicarakan di jejaring media sosial terutama *Twitter*. Salah satunya topik terkait Penerapan Permendikbud Nomor 30 Tahun 2021. Pada *tweet* dapat dilakukan analisis sentimen untuk mengetahui tanggapan masyarakat terhadap penerapan Permendikbud PPKS dengan melakukan identifikasi dan mengkategorikan polaritas teks untuk menentukan text tersebut memiliki nilai positif atau negatif. Model sistem bertujuan sebagai evaluasi pihak pemerintah dalam tindak lanjut dari penetapan kebijakan. Pengumpulan data yang diambil dari *twitter* dengan *hash tag* kesetaraan gender, kekerasan seksual, dan PPKS berjumlah 159 data. Pembagian data latih dan data uji menggunakan rasio 70:30 secara acak oleh sistem. Untuk data latih digunakan sebanyak 111 data, sedangkan data uji digunakan sebanyak 48 data. Berdasarkan hasil pengujian dengan *textblob* yang dibantu dengan menggunakan algoritma *Support Vector Machine* menggunakan data uji sebanyak 48 data dengan 34 data positif, dan 14 data negatif, mendapatkan akurasi sebesar 70,8%.

ABSTRACT

In an all-digital era, all aspects of implementing a policy by the government are topics that are widely discussed on social media, especially Twitter. One of them is a topic related to the Implementation of Permendikbud Number 30 of 2021. In the tweet, a sentiment analysis can be carried out to find out the public's response to the application of the PPKS Permendikbud by identifying and categorizing the polarity of the text to determine that the text has a positive or negative value. The system model aims to evaluate the government in the follow-up of policy determination. Data collection taken from twitter with hash tags of gender equality, sexual violence, and PPKS amounted to 159 data. The division of training data and test data uses a random 70:30 ratio by the system. For training data, 111 data were used, while test data was used as much as 48 data. Based on the test results with textblob assisted by using the Support Vector Machine algorithm using test data as many as 48 data with 34 positive data, and 14 negative data, it obtained an accuracy of 70.8%.

Corresponding Author :

Yesy Afrillia

Program Studi Teknik Informatika, Jurusan Informatika, Fakultas Teknik, Universitas Malikussaleh
Kampus Bukit indah, Jl. Batam, Blang Pulo, Kecamatan Muara Satu, Kota Lhokseumawe, Aceh, 24352

Email: yesy.afrillia@unimal.ac.id

PENDAHULUAN

Kesetaraan gender merupakan isu yang hangat diperbincangkan dalam ruang lingkup sosial saat ini, Makna kesetaraan gender bukanlah penerapan dalam ranah kesetaraan yang menuntut pemenuhan standarisasi antara kaum laki-laki dan perempuan melainkan penempatan kedudukan yang sama dalam perpektif hakekatnya dalam menyuarkan aspirasi.

Kekerasan seksual merupakan bentuk tindakan kejahatan yang dapat dilakukan kapan saja dan ditempat mana saja, termasuk diruang lingkup pendidikan seperti di kampus. Pada akhir-akhir ini atau dalam beberapa tahun terakhir pemberitaan media terhadap kekerasan seksual lebih tertuju ke ranah perguruan tinggi.

Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek) memiliki komitmen dalam mengeradikasi 'tiga dosa besar' dunia pendidikan, yaitu perundungan, intoleransi, dan kekerasan seksual. Segala upaya telah ditempuh, yang saat ini yaitu penetapan Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi (Permendikbudristek) Nomor 30 Tahun 2021 tentang Pencegahan dan Penanganan Kekerasan Seksual di lingkungan Perguruan Tinggi (Permendikbudristek PPKS). Permendikbudristek PPKS ini hadir menjadi solusi atas terjadinya berbagai kasus kekerasan seksual yang terjadi di lingkungan perguruan tinggi (Kemdikbud, 2021).

Permendikbudristek tentang PPKS dirancang untuk menerapkan moral serta akhlak mulia sebagai maksud utama dalam pendidikan di Indonesia. Sebagaimana yang sudah termuat dalam UUD, UU No 20 Tahun 2003, UU No 12 Tahun 2012, dan Permendikbud No 3 Tahun 2020. Terbitnya peraturan menteri ini tentunya memiliki respon pro dan kontra dari kalangan masyarakat secara luas. Serta dukungan terhadap penetapan Permendikbud Ristek Nomor 30 Tahun 2021 terkait kegiatan Pencegahan dan Penanganan Kekerasan Seksual (PPKS) sebagai implementasi di Lingkungan Perguruan Tinggi juga banyak diberikan masyarakat melalui unggahan berupa komentar di media sosial.

Di era yang serba digital, segala aspek penerapan suatu kebijakan oleh pemerintah menjadi topik yang banyak dibicarakan di jejaring sosial terutama Twitter. Twitter merupakan media sosial yang menghubungkan user dengan berbagai komentar, opini, pandangan, cerita, ide dan berita terbaru mengenai hal yang dianggap menarik oleh banyak orang. Twitter merupakan media sosial (medsos) yang datanya cukup banyak digunakan untuk beberapa penelitian sosial seperti opini publik (Puspita, 2020).

Pada tweet dapat dilakukan analisis sentimen untuk mengetahui respon dari masyarakat terhadap penerapan Permendikbud Nomor 30 Tahun 2021 dalam melakukan identifikasi dan mengkategorikan polaritas sebuah teks untuk menentukan apakah dokumen tertentu memiliki nilai positif atau negatif yang sesuai dengan kategori yang telah ditentukan.

Salah satu metode yang dapat digunakan untuk melakukan analisis sentimen yaitu dengan menerapkan *Lexicon Based* yaitu menggunakan *Textblob*. *Textblob* merupakan salah satu *library* pada Python yang bisa digunakan untuk memproses data secara tekstual (Hernikawati, 2021).

METODE PENELITIAN

A. Pengumpulan Data

Melakukan koordinasi serta penerapan kriteria data yang dikumpulkan yang dilakukan pada Dinas Pemberdayaan Perempuan dan Anak Kabupaten Aceh Utara dan Kota Lhokseumawe. Dengan mendapatkan hasil yaitu pengambilan komentar pada twitter dengan kriteria data yaitu : #Permendikbudristek30, #PPKS, #SAHKANRUUPKS, ruupks dan tpks. Data yang digunakan merupakan hasil *scraping* dari komentar-komentar pengguna *twitter* mulai tahun 2021-2022 terkait isu kesetaraan gender dalam penerapan permendikbudristek nomor 30 tahun 2021 yang telah diberi label. Keterkaitan antara *hagtag* yang dipilih dengan Permendikbudristek Nomor 30 Tahun 2021 ialah hal yang sering dan banyak digunakan para pengguna *twitter* dalam memberikan komentar di media social *twitter* untuk pembahasan kebijakan Permendikbudristek Nomor 30 Tahun 2021.

B. Kajian Pustaka

Kajian Pustaka yang dilakukan yaitu mengumpulkan dan mempelajari buku literatur atau jurnal serta bacaan lainnya yang berhubungan dengan sistem analisis sentimen, sistem data mining khususnya text mining serta metode yang digunakan pada sistem yaitu metode textblob dan SVM. Tahapan ini dilakukan untuk memperoleh data sekunder pada penelitian.

Keseluruhan data yang sudah dilakukan *textpreprocessing* akan dibagi menjadi data *training* dan data *testing* dengan rasio data 70:30. Dimana keseluruhan data yang digunakan ini yaitu 159 data. Kemudian dibagi untuk data latih 111 data dan 48 untuk data uji. Jumlah data uji yang digunakan 48 data dimana 34 data positif, 14 data negatif. Dan didapatkan nilai Presisi sebesar 35,4%, Recall sebesar 50%, dan F1-Score sebesar 41,4% dan nilai tingkat akurasi sebesar 70,8%.

C. Analisa Sistem

Selama ini untuk mempelajari tingkat sentimen masyarakat dalam suatu penerapan kebijakan hanya dilakukan pengumpulan data dari suara perwakilan masyarakat saja misalnya pembahasan pada suatu forum DPRK, DPRA hingga pada level atas yaitu Dewan Perwakilan Rakyat yang mana pihak-pihak tersebut juga terlibat langsung dalam rancangan penerapan dari suatu kebijakan tersebut. Dengan adanya teknologi saat ini maka untuk mengetahui tingkat sentimen masyarakat dalam penerapan suatu kebijakan tersebut kita bisa langsung mengambil data valid dari komentar langsung masyarakat melalui sosial media salah satunya yaitu Twitter.

D. Implementasi dan Pengujian

Setelah data dimasukkan dalam database maka tahap selanjutnya ialah mengimplementasikan dalam melakukan pengujian model sistem untuk menarik informasi yang dibutuhkan peneliti yaitu mengetahui tingkat sentimen masyarakat terhadap penerapan Permendikbudristek Nomor 30 Tahun 2021.

E. Support Vector Machine

Support Vector Machine (SVM) adalah algoritma pembelajaran untuk menganalisis data serta memperkenalkan pola yang digunakan untuk klasifikasi. SVM juga merupakan salah satu model dari machine learning dari turunan teori statistik untuk memberikan hasil yang lebih baik dalam penentuan tingkat keakuratan dibandingkan dengan metode *machine learning* lainnya.

Dalam metode SVM, objek informasi terluar yang paling dekat dengan *hyperplane* disebut *support vector*. Hanya saja *support vector* ini yang diperhitungkan oleh SVM untuk menemukan *hyperplane* yang paling ideal sedangkan objek informasi lainnya tidak diperhitungkan sama sekali (Irawan, 2018).

Proses pertama yang dilakukan adalah mengimport library pandas untuk membaca data dan library Sklearn untuk memanggil algoritma Support Vector Machine selanjutnya yang dilakukan adalah membaca data dari dataset berformat .csv kemudian membagi data menjadi dua bagian yaitu data latih dan data uji. Selanjutnya data yang telah dibagi akan diuji menggunakan algoritma Support Vector Machine untuk mendapatkan hasil pengujian terhadap sentiment polarity yang didapat menggunakan analisis textblob.

F. Confusion Matriks

Confusion Matrix adalah suatu proses yang dipakai dalam menganalisa keakuratan dari model klasifikasi yang dibuat untuk mengidentifikasi data dengan kelas yang berbeda. Dengan diukurinya tingkat keakuratan maka dapat diketahui performa dari suatu model klasifikasi tersebut. Hasil akurasi dari sebuah klasifikasi dapat diketahui dari hasil data latih dengan bentuk persentase dari kelompok data latih, yang merupakan hasil dari klasifikasi yang benar (Nugroho, 2019). Sedangkan, hasil uji dari model klasifikasi yang akan dianalisa menggunakan tabel atau sering disebut (*confusion matrix*) dapat memperoleh nilai *accuracy*, *recall*, *precision*, serta *f1-score*. Berikut *Confusion matrix* yang dievaluasi dengan tabel yang menerangkan bahwa hasil klasifikasi dari jumlah data testing yang benar dan dari jumlah data testing yang salah (Anggreany, 2022).

Tabel 1. *Confusion Matrix*

	Data Prediksi		
	Positif	Negatif	Netral
Positif	TP	FP	PN
Negatif	FN	TNg	NgN
Netral	NP	NNg	TN

Keterangan :

- TP = Prediksi Benar Bernilai Positif
- FN = Negatif Terprediksi Positif
- NP = Netral Terprediksi Positif
- FP = Positif Terprediksi Negative
- TNg = Prediksi Benar Bernilai Negatif
- NNg = Netral Terprediksi Negatif
- PN = Positif Terprediksi Netral
- NgN = Negatif Terprediksi Netral
- TN = Terprediksi Netral

Nilai akurasi adalah persentase terhadap jumlah *record* data yang benar diklasifikasikan oleh algoritma sesudah pengujian dibagi dengan jumlah total *record*. *Precision* adalah jumlah komentar positif yang diprediksi sebagai nilai positif pada data aktual. *Recall* adalah jumlah komentar positif aktual yang diprediksi sebagai nilai positif. Untuk menghitung tingkat keakurasian, *precision recall* dan *f1score* dapat dilihat pada persamaan berikut :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

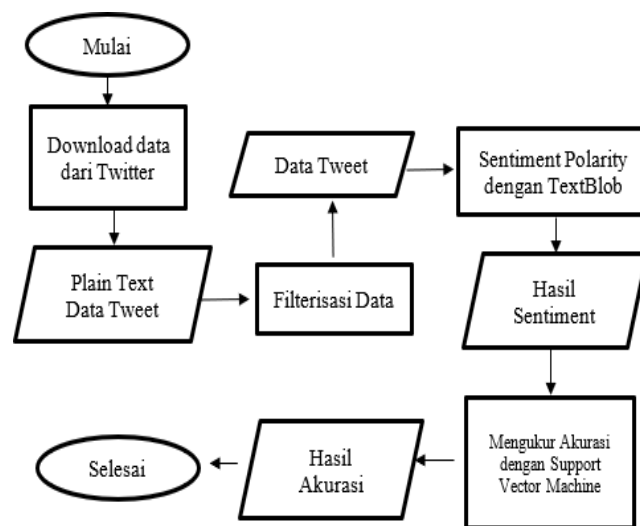
$$Recall = \frac{TP}{TP+FN} \tag{2}$$

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \tag{4}$$

G. Skema Sistem

Skema sistem yang terbentuk adalah sebagai berikut :



Gambar 1. Skema sistem

Flowchart ini menggambarkan skema model sistem yang dibangun untuk mendapatkan hasil kategori sentiment suatu text. Yang mana dimulai dari pengambilan data berupa komentar

dari tweet menggunakan teknik scraping, kemudian disimpan ke database sistem. Setelah itu data akan dipisah menjadi data training dan data testing atau sering disebut Split Data. kemudian text tweet masuk dalam filterisasi data, yaitu tahapan dari pre-processing untuk mendapatkan teks yang diperlukan dan menghilangkan teks yang tidak penting. kemudian mendapatkan data bersih kemudian dibangun model menggunakan textblob. Kemudian masuk pada tahapan uji coba sistem, yaitu kegiatan untuk menganalisis tingkat sentiment pengguna twitter sesuai dengan studi kasus yang dibahas yaitu komentar pengguna twitter terhadap penerapan Permendikbud Nomor 30 tahun 2021. Dengan penerapan model ini maka akan mendapatkan hasil berupa informasi klasifikasi komentar/opini pengguna twitter dalam pembahasan penerapan Permendikbud Nomor 30 tahun 2021, Dengan pernyataan hasil berupa persentase pengguna twitter yang masuk dalam kategori Pro atau Kontra terhadap isu tersebut. Pro yang dimaksudkan masuk sebagai kategori opini positif dan Kontra sebagai kategori penyampaian opini negatif oleh pengguna twitter. Hal ini dilakukan secara otomatis oleh model yang telah dibangun.

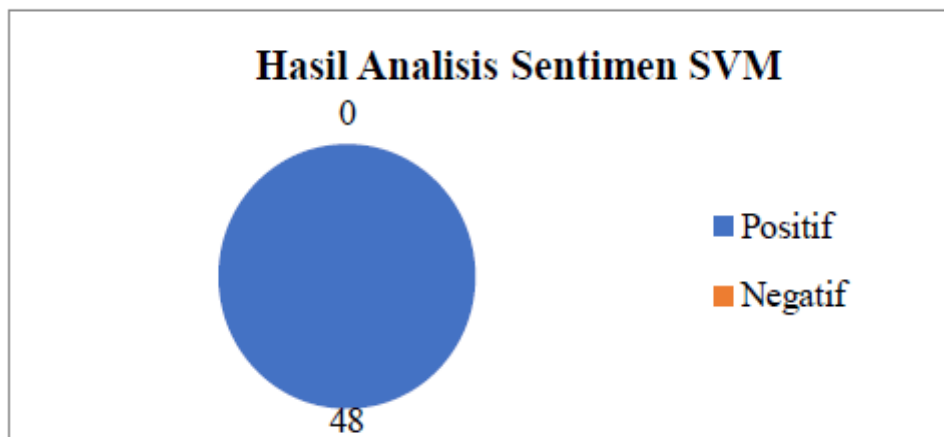
HASIL DAN PEMBAHASAN

Data *training* yang digunakan yaitu 111 ulasan dan data *tasting* yang digunakan sebanyak 48 ulasan. Berdasarkan data *tasting* tersebut, nilai aktual positif yang diuji sebanyak 34 data, aktual negatif 14 data, sehingga didapatkan *confusion matrix* pada tabel 2. sebagai berikut :

Tabel 2. Confusion Matrix data testing

		Prediksi	
		Positif	Negatif
Aktual	Positif	34	0
	Negatif	14	0

Aktual merupakan nilai asli dari label/*class*. Sedangkan prediksi merupakan nilai prediksi hasil pemodelan *machine learning*. Dari tabel tersebut dibentuklah grafik sebagai berikut:



Gambar 2. Hasil Analisis Sentimen SVM

Dari grafik di atas dapat disimpulkan bahwa:

- Dari jumlah positif 34 data menghasilkan prediksi 48 data positif.
- Dari jumlah negatif 14 data menghasilkan prediksi 0 data negatif.

Kemudian dihitung akurasi, presisi, *recall*, dan *f1-score* untuk menentukan hasil *performance score* dari metode *Support Vector Machine* berdasarkan tabel 2 yang dijadikan sebagai validitas dari kinerja klasifikasi algoritma *Support Vector Machine*.

Berikut perhitungannya :

1. Perhitungan Akurasi sebagai berikut :

$$Akurasi = \frac{34 + 0}{34 + 14} \times 100\% = \frac{34}{48} \times 100\% = 70,8\%$$

2. Perhitungan *precision* tiap-tiap kelas sebagai berikut :

$$Presisi (positif) = \frac{TP}{TP + FN_g} = \frac{34}{34 + 14} = \frac{34}{48} = 0,708$$

$$Presisi (negatif) = \frac{TN_g}{FN_g + FP} = \frac{0}{0 + 0} = \frac{0}{0} = 0$$

Maka hasil perhitungan presisi keseluruhan sebagai berikut :

$$\begin{aligned} Presisi &= \frac{Presisi(pos) + Presisi(neg)}{2} \times 100\% \\ &= \frac{0,708 + 0}{2} \times 100\% = 35,4\% \end{aligned}$$

3. Perhitungan *recall* tiap-tiap kelas sebagai berikut:

$$Recall (positif) = \frac{TP}{TP + FP} = \frac{34}{34 + 0} = \frac{34}{34} = 1$$

$$Recall (negatif) = \frac{TN_g}{TN_g + FN} = \frac{0}{0 + 14} = \frac{0}{14} = 0$$

Maka hasil perhitungan recall keseluruhan tiap kelas sebagai berikut :

$$\begin{aligned} Recall &= \frac{Recall(pos) + Recall(neg)}{2} \times 100\% \\ &= \frac{1 + 0}{2} \times 100\% = 50\% \end{aligned}$$

4. Perhitungan *F1-Score* tiap-tiap kelas sebagai berikut :

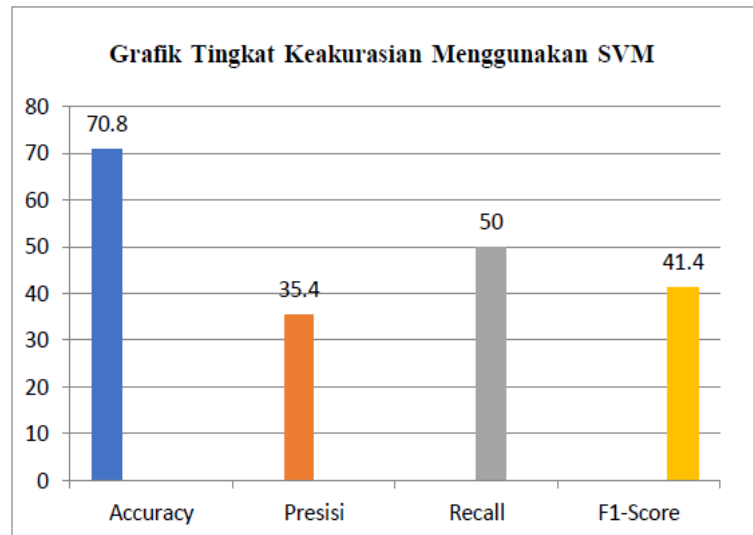
$$\begin{aligned} F1 - Score (positif) &= 2 \times \frac{Presisi (positif) \times Recall (positif)}{Presisi (positif) + Recall(positif)} \\ &= 2 \times \frac{0,708 \times 1}{0,708 + 1} = 2 \times \frac{0,708}{1,708} \\ &= 2 \times 0,4145 = 0,829 \end{aligned}$$

$$\begin{aligned} F1 - Score (negatif) &= 2 \times \frac{Presisi (negatif) \times Recall (negatif)}{Presisi (negatif) + Recall(negatif)} \\ &= 2 \times \frac{0 \times 0}{0 + 0} = 2 \times 0 = 0 \end{aligned}$$

Maka perhitungan *F1-Score* keseluruhannya sebagai berikut :

$$\begin{aligned} F1 - Score &= \frac{F1 - Score(pos) \times F1 - score (neg)}{2} \times 100\% \\ &= \frac{0,829 + 0}{2} \times 100\% = \frac{0,829}{2} \times 100\% = 41,4\% \end{aligned}$$

Dari perhitungan *confusion matrix* di atas dibuat grafik tingkat keakurasian menggunakan SVM sebagai berikut :



Gambar 3. Tingkat Keakurasian Menggunakan SVM

Dari grafik di atas didapatkan nilai tingkat akurasi sebesar 70,8%, *Presisi* sebesar 35,4%, *Recall* sebesar 50%, dan *F1-Score* sebesar 41,4%.

KESIMPULAN DAN SARAN

Kesimpulan

Model sistem ini dapat memberikan manfaat dalam menganalisis tingkat sentimen masyarakat tentang penerapan Permendikbudristek Nomor 30 Tahun 2021 yaitu tentang PPKS sebagai evaluasi pihak pemerintah dalam tindak lanjut dari penetapan kebijakan dari pengumpulan data yang di ambil dari tweet dengan hashtag kesetaraan gender, kekerasan seksual, dan PPKS berjumlah 159 data, kemudian didapatkan 121 tweet dari keseluruhan yang mengandung komentar positif artinya meberikan opini yang bagus terhadap Permendikbudristek Nomor 30 Tahun 2021 tentang Pencegahan dan Penanganan Kekerasan Seksual di lingkungan Perguruan Tinggi (Permendikbudristek PPKS). 38 tweet dari keseluruhan tweet mengandung sentimen negatif artinya memberikan opini tidakbaik kepada pemerintah terhadap Pencegahan dan Penanganan Kekerasan Seksual di lingkungan Perguruan Tinggi (Permendikbudristek PPKS). Pembagian data latih dan data uji dilakukan menggunakan rasio 70:30 secara acak oleh sistem. Dimana data latih sebanyak 111 data, data uji sebanyak 48 data. Berdasarkan hasil pengujian dengan *textblob* dibantu dengan menerapkan algoritma *Support Vector Machine* menggunakan data uji sebanyak 48 data dengan 34 data positif, dan 14 data negatif, penelitian ini mendapatkan akurasi sebesar 70,8%.

Saran

Diharapkan penelitian ini dapat dilanjutkan dengan menggunakan subjek dan metode yang lain untuk dapat membandingkan hasil keakuratan sistem dari masing-masing metode tersebut.

UCAPAN TERIMA KASIH

Terimakasih kepada Stakeholder yang terlibat dalam pengambilan kriteria data latih dan data uji yaitu Dinas Pemberdayaan Perempuan Kabupaten Aceh Utara dan Kota Lhokseumawe.

REFERENSI

Anggreany, M. S. (2022). Confusion Matrix. *socs.binus.ac.id*.
<https://socs.binus.ac.id/2020/11/01/confusion-matrix/> (accessed Sep. 02, 2022).

- Hernikawati, D. (2021). Kecenderungan Tanggapan Masyarakat Terhadap Vaksin Sinovac Berdasarkan Lexicon Based Sentiment Analysis (The Trend of Public Response to Sinovac Vaccine Based on Lexicon Based Sentiment Analysis). *Jurnal Iptekom (Jurnal Ilmu Pengetahuan & Teknologi Informasi)*, 23(1), 21-31.
- Irawan, F., & Samopa, F. (2018, July). A Comparative Assessment of the Random Forest and SVM Algorithms Using Combination of Principal Component Analysis and SMOTE For Accounts Receivable Seamless Prediction case study company X in Surabaya. *In Proceedings of the 2nd International Seminar Of Contemporary ONBUSINESS Research & Management. University Warmadewa Bali*.
- Kemendikbud dan Kebudayaan. (2021). Merdeka Belajar Episode Keempat Belas Tetaskan Solusi untuk Kekerasan Seksual di Perguruan Tinggi. *kemdikbud.go.id*, 2021. <https://www.kemdikbud.go.id/main/blog/2021/11/merdeka-belajar-episode-keempat-belas-tetaskan-solusi-untuk-kekerasan-seksual-di-perguruan-tinggi> (accessed Sep. 03, 2022).
- Nugroho, K. S. (2019). Confusion matrix untuk evaluasi model pada supervised learning. Diakses dari: <https://ksnugroho.medium.com/confusion-matrix-untuk-evaluasi-model-pada-supervised-machine-learning-bc4b1ae9ae3f>.
- Puspita, R., & Suciati, T. N. (2020). Mobile Phone dan Media Sosial: Penggunaan dan Tantangannya pada Jurnalisme Online Indonesia. *Ekspresi Dan Persepsi: Jurnal Ilmu Komunikasi*, 3(2), 132-146.. doi: 10.33822/jep.v3i2.1781.
- Suherman, A., Aryani, L., & Yulyana, E. (2021). Analisis Fungsi Peraturan Menteri Pendidikan, Kebudayaan, Riset dan Teknologi Nomor 30 Tahun 2021 dalam Mencegah Kekerasan Seksual di Kampus. *Jurnal Ilmiah Wahana Pendidikan*, 7(7), 173-182., doi: 10.5281/zenodo.5704133.
- Wandani, A., Fauziah, F., & Andrianingsih, A. (2021). Sentimen Analisis Pengguna Twitter pada Event Flash Sale Menggunakan Algoritma K-NN, Random Forest, dan Naive Bayes. *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, 5(2), 651-665. Available: <https://tunasbangsa.ac.id/ejurnal/index.php/jsakti>