

## ANALISIS KERNEL DAN PARAMETER *COST* PADA KLASIFIKASI CUACA HARIAN MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE*

Dyyah Qurniatun\*, Sitta Alief Farihati

Program Studi Matematika, Universitas Terbuka, Tangerang Selatan, Indonesia

\*Penulis korespondensi: [dyyqrn@gmail.com](mailto:dyyqrn@gmail.com)

### ABSTRAK

Salah satu dampak posisi Indonesia sebagai negara tropis adalah mudahnya cuaca berubah dari berawan menjadi hujan lebat. Prediksi cuaca dibutuhkan sebagai alat peringatan dini yang bermanfaat di berbagai sektor. Untuk menangani permasalahan tersebut diperlukan prediksi prakiraan cuaca yang akurat. Salah satu model yang dapat digunakan adalah *Support Vector Machine* (SVM). Artikel ini akan mengkaji model algoritma klasifikasi SVM dengan kernel dan parameter *cost C*. Kernel yang akan dikaji adalah *linear*, *Radial Basis Function* (RBF), *polinomial*, dan *sigmoid*. Nilai *C* yang akan digunakan adalah 1, 10, 100, dan 1000. *Output* prediksi cuaca dalam kasus ini adalah kategori berawan atau kategori hujan. Variabel *input* yang digunakan adalah temperatur, kelembaban, curah hujan, lama penyinaran matahari, dan kecepatan angin. Data observasi yang digunakan adalah data harian periode 1 Januari 2016 – 31 Desember 2023 dari Stasiun Geofisika Sleman. Perbandingan data latih dan data uji yang digunakan sebesar 80:20. Hasil yang diperoleh menunjukkan bahwa parameter untuk model klasifikasi SVM yang memberikan akurasi terbaik adalah kernel *linear* dengan nilai *C* 10 dengan akurasi sebesar 100%.

**Kata kunci:** *cost*, cuaca harian, kernel, klasifikasi, SVM

### 1 PENDAHULUAN

Indonesia merupakan negara tropis dengan keanekaragaman topografi yang meliputi daerah pegunungan, dataran rendah, dan pantai yang banyak. Letak geografis dan astronomis Indonesia secara signifikan mempengaruhi kondisi iklim yang beragam di setiap daerahnya. Keragaman iklim ini dipengaruhi oleh aktivitas iklim yaitu fenomena global *El Nino Southern Oscillation* (ENSO) dan *Indian Ocean Dipole* (IOD), fenomena regional sirkulasi angin Monsun Asia-Australia, daerah pertemuan angin antar tropis atau *Inter Tropical Convergence Zone* (ITCZ), dan suhu permukaan laut di sekitar wilayah Indonesia (BMKG, 2024).

Aktivitas iklim yang cepat berubah berdampak terhadap cuaca di berbagai wilayah terutama di daerah yang rentan terhadap perubahan iklim. BMKG merupakan lembaga pemerintah yang melakukan tugas pelaksanaan, pembinaan dan pengendalian observasi, dan pengolahan data dan informasi di bidang meteorologi, klimatologi, dan geofisika di Indonesia. Pengamatan iklim harian dilakukan di stasiun/UPT (Unit Pelaksana Teknis) yang tersebar di seluruh Indonesia. Adapun parameter yang diamati diantaranya ialah kelembapan udara, curah hujan, temperatur, kecepatan angin, arah angin, dan lama penyinaran matahari (Sunarmi dkk., 2022). Observasi cuaca telah lama dilakukan di berbagai daerah menggunakan bermacam model penelitian dan menghasilkan data yang panjang. Data pengamatan ini kemudian digunakan untuk keperluan penelitian, salah satunya adalah pemanfaatan data curah hujan untuk membuat model prediksi yang akurat (Jayadianti, 2020). Data prediksi akumulasi curah hujan selama 24 jam di Indonesia tersedia di website BMKG, dengan data model prediksi cuaca berbentuk numerik.

Salah satu metode untuk mendapatkan hasil prediksi yang akurat dan efisien adalah *machine learning*. *Machine learning* merupakan bagian dari algoritma kecerdasan buatan yang digunakan secara luas untuk menyelesaikan masalah optimasi yang kompleks. Algoritma *machine learning* secara otomatis dapat melatih dan menganalisis suatu dataset untuk menghasilkan model yang akurat (Mardiansyah dkk., 2022). *Support Vector Machine* (SVM) adalah salah satu model *machine learning* yang mampu melakukan klasifikasi *linear* ataupun *non-linear*. SVM bekerja dengan memisahkan dua himpunan data dari dua kelas yang berbeda dengan memaksimalkan batas fungsi pemisah atau *hyperplane* (Dasmasele dkk., 2022). Model klasifikasi ini dapat digunakan untuk melakukan prediksi kategori cuaca berawan atau hujan.

Pada tahun 2024, Pridiptama melakukan penelitian tentang perbandingan algoritma SVM dan Naive Bayes pada klasifikasi penyakit tekanan darah tinggi. Salah satu hasil dari penelitian tersebut adalah keakuratan SVM lebih baik dibanding algoritma Naive Bayes. Adapun penelitian dengan membandingkan tiga jenis kernel yang berbeda dengan berbagai nilai parameter *cost C* telah dilakukan oleh Pratiwi dan Setyawan pada tahun 2021. Hasilnya adalah akurasi terbaik diperoleh pada fungsi kernel linier dan polinom dengan nilai *C* 1 dan 1000 sebesar 78,38%. Penelitian serupa juga telah dilakukan oleh Dasmasele yang menghasilkan akurasi sebesar 95% pada penggunaan kernel *Radial Basis Function* (RBF) untuk model SVM non-linier dengan pembagian data 80:20.

Berdasarkan uraian yang telah dijelaskan, dalam artikel ini akan ditunjukkan penggunaan model klasifikasi SVM dalam penentuan kategori cuaca di Kabupaten Sleman. Dalam hal ini, akan digunakan empat metode kernel yaitu *linear*, RBF, polinomial, dan sigmoid, dengan nilai *C* 1, 10, 100, dan 1000.

## 2 METODE

Data yang digunakan dalam artikel ini adalah dataset dari data iklim yang bersifat *open source* yang didapatkan dari website Data Online – Pusat Database – BMKG dengan alamat [dataonline.bmkg.go.id/](http://dataonline.bmkg.go.id/). Data observasi merupakan data harian periode 1 Januari 2016 – 31 Desember 2023 dari Stasiun Geofisika Sleman dengan lokasi observasi lintang -7.82000, bujur 110.30000, dan elevasi 153.

### 2.1 Pemrosesan Data

Data iklim periode 1 Januari 2016 – 31 Desember 2023 digabungkan menggunakan *Microsoft Excel* kemudian disimpan dalam format *csv*. Dataset ini kemudian diolah menggunakan bahasa pemrograman *Python*. BMKG mengkategorikan intensitas hujan sebanyak enam kategori seperti yang tercantum dalam Tabel 1.

**Tabel 1.** Kategori Intensitas Hujan

Intensitas Hujan	Jumlah Curah Hujan
Berawan	0 mm/hari
Hujan Ringan	0,5 – 20 mm/hari
Hujan Sedang	20 – 50 mm/hari
Hujan Lebat	50 – 100 mm/hari
Hujan Sangat Lebat	100 – 150 mm/hari
Hujan Ekstrem	lebih dari 150 mm/hari

Dalam artikel ini akan dilakukan penggolongan menjadi dua kategori yaitu berawan dan hujan. Jumlah curah hujan kategori berawan sebesar 0 mm/hari, sedangkan kategori hujan lebih dari 0,5 mm/hari. Adapun variabel yang digunakan dalam klasifikasi SVM dapat dilihat pada Tabel 2. Variabel X1 hingga variabel X8 merupakan variabel masukan dan variabel Y adalah variabel keluaran. Variabel keluaran Y terdapat dua kelas yaitu berawan (0) dan hujan (1).

**Tabel 2.** Variabel

Variabel	Keterangan
X1	Temperatur Minimum
X2	Temperatur Maksimum
X3	Temperature Rata-Rata
X4	Kelembaban Rata-Rata
X5	Curah Hujan
X6	Lama Penyinaran Matahari
X7	Kecepatan Angin Maksimum
X8	Kecepatan Angin Rata-Rata
Y	Berawan (0), Hujan (1)

Proses selanjutnya adalah pembagian data *train/test* yang digunakan untuk melatih dan menguji data. Dalam hal ini, perbandingan data latih dan data uji yang digunakan adalah sebesar 80:20.

## 2.2 Support Vector Machine

*Support Vector Machine* (SVM) pertama kali dikenalkan oleh Vapnik, merupakan algoritma *machine learning* yang dapat digunakan dalam penyelesaian masalah klasifikasi, regresi dan deteksi (Bishop, 2006). SVM bekerja dengan cara mengklasifikasi perbedaan kelas antara satu kelas dengan kelas lainnya melalui sebuah *hyperplane* multi-dimensi (Lombay & Hernandez, 2021). *Hyperplane* adalah pemisah terbaik antara dua kelas yang didapatkan dengan cara mengukur margin *hyperplane*, yang merupakan data terdekat dari masing-masing kelas, yang kemudian dapat dicari titik maksimalnya (Putra dkk., 2020).

Menurut James (2023), SVM adalah pengembangan dari *support vector classifier* yang dihasilkan dengan memperbesar ruang fitur secara khusus menggunakan kernel. Misalkan diberikan data latih  $(x_i, y_i)$  dengan  $x_i$  adalah observasi ke- $i$  dan  $y_i$  adalah label dari observasi ke- $i$ . Kemudian untuk menentukan *hyperplane*, yang dibutuhkan untuk memaksimalkan margin dua kelas digunakan.

$$w^T x + b = 0 \quad (1)$$

dengan  $w$  adalah vektor bobot dan  $b$  adalah bias.

Margin adalah jarak antara *hyperplane* dan titik data terdekat dari dua kelas. Jarak *hyperplane* untuk sebuah titik  $x_i$  adalah:

$$\frac{|w^T x_i + b|}{\|w\|} \quad (2)$$

dengan meminimalkan  $\|w\|$ , margin dapat dimaksimalkan sehingga data dapat terklasifikasi dengan benar.

Kasus ideal memberikan,

$$y_i(w^T x_i + b) \geq 1 \quad (3)$$

dengan formulasi primal, fungsi objektif menjadi:

$$\min_{w,b} \frac{1}{2} \|w\|^2 \quad (4)$$

dengan batasan:

$$y_i(w^T x_i + b) \geq 1, \forall i \quad (5)$$

Data yang tidak dapat dipisahkan secara *linear*, digunakan variable *slack*  $\xi_i \geq 0$ :

$$y_i(w^T x_i + b) \geq 1 - \xi_i, \forall i \quad (6)$$

Fungsi objektif kemudian menjadi:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \quad (7)$$

dengan  $C$  adalah parameter regulasi yang mengontrol dalam pembuatan keputusan antara margin yang lebar dan pelanggaran batas margin.

Masalah fungsi objektif dapat diselesaikan dengan mengkonversi ke bentuk dual:

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (8)$$

dengan batasan:

$$0 \leq \alpha_i \leq C, \sum_{i=1}^n \alpha_i y_i = 0 \quad (9)$$

$\alpha_i$  adalah variabel dual yang berhubungan dengan setiap data latih.

Setelah menyelesaikan masalah optimasi dan memperoleh nilai  $\alpha_i$ , prediksi untuk data baru  $x$  dilakukan dengan:

$$f(x) = \text{sgn} \left( \sum_{j=1}^n \alpha_j y_j x_j^T x + b \right) \quad (10)$$

Dalam SVM, terdapat kernel *trick* yang dapat digunakan dalam klasifikasi untuk mendapatkan hasil yang lebih optimal (Fremmuzar & Baita, 2023). Menurut Awad dan Khanna (2015), beberapa fungsi kernel yang populer adalah fungsi kernel *linear*, RBF, fungsi *polinomial*, dan *hyperbolic tangent* (sigmoid).

Kernel untuk SVM *linear* adalah:

$$K(x_i, x_j) = x_i^T x_j \quad (11)$$

Untuk SVM non-*linear*, fungsi kernel yang digunakan adalah:

1. Kernel RBF:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (12)$$

2. Kernel Polinomial:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (13)$$

3. Kernel Sigmoid:

$$K(x_i, x_j) = \tanh(kx_i^T x_j + \theta) \quad (14)$$

Dengan demikian, untuk prediksi data baru pada SVM non-linear digunakan:

$$f(x) = \text{sgn} \left( \sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \right) \quad (15)$$

### 2.3 Evaluasi Model

Sebelum model prediksi dikembangkan, semua model harus dievaluasi menggunakan berbagai parameter evaluasi (Belete & Huchaiyah, 2021). Evaluasi model dilakukan menggunakan *confusion matrix* kemudian dihitung nilai akurasi, *precision*, *recall*, dan *F1-Score*. Tabel 3 memuat penjelasan mengenai *confusion matrix* sebagai berikut:

**Tabel 3.** *Confusion matrix*

Nilai Prediksi	Nilai Aktual	Tipe	Keterangan
1	1	True Positive (TP)	Prediksi benar dan itu benar
0	0	True Negative (TN)	Prediksi salah dan itu benar
1	0	False Positive (FP)	Prediksi benar dan itu salah
0	1	False Negative (FN)	Prediksi salah dan itu salah

Akurasi dihitung sebagai berikut:

$$\text{Akurasi} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad (16)$$

*Precision* dihitung sebagai berikut:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (17)$$

*Recall* dihitung sebagai berikut:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (18)$$

*F1 - Score* dihitung sebagai berikut:

$$F1 - \text{Score} = 2 \times \left( \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right) \quad (19)$$

## 3 HASIL DAN PEMBAHASAN

Data iklim yang diambil sebanyak 2922 data, kemudian dilakukan pemrosesan data untuk menghilangkan data tidak terukur dan tidak ada data. Data bersih yang didapatkan adalah

sebanyak 2193 data. Gambar 1 merupakan data lima baris pertama dari 2193 data bersih yang akan dilanjutkan ke proses selanjutnya.

Dataset memiliki *header* yang berisi variabel-variabel. Kolom X1, X2, X3 secara berturut-turut merupakan temperatur minimum, maksimum dan rata-rata dengan satuan derajat celsius ( $^{\circ}\text{C}$ ). X4 adalah kelembaban rata-rata dalam %, kolom X5 adalah curah hujan dalam mm/hari, dan X6 adalah lama penyinaran sinar matahari dalam hitungan jam. Selanjutnya, kolom X8 dan X9 merupakan kecepatan angin rata-rata dan kecepatan angin maksimum dalam satuan knot. Kolom Y adalah kategori cuaca yang memiliki dua kelas yaitu berawan (0) dan hujan (1). Tipe data seluruh variabel X adalah *float* sedangkan variabel Y memiliki tipe data *integer*.

x1	x2	x3	x4	x5	x6	x7	x8	Y
25.0	31.4	27.6	83.0	0.0	8.2	2.0	0.0	0
24.0	33.3	28.2	80.0	0.0	6.9	2.0	0.0	0
25.0	32.0	27.4	88.0	0.0	9.4	2.0	1.0	0
23.0	34.4	27.7	82.0	4.4	6.1	2.0	1.0	1
24.0	33.8	27.6	87.0	0.0	8.7	2.0	0.0	0

**Gambar 1.** Contoh dataset yang sudah bersih

Dari 2193 data, terhitung terdapat 1137 data berawan (0) dan 1056 data hujan (1). Dengan demikian, persentase data berawan dibanding data hujan adalah 51,85% dan 48,15%. Pembagian data *training* dan data *testing* adalah 80:20. Data latih sebanyak 1754 dan data uji sebanyak 439 data.

### 3.1 Klasifikasi SVM

Hasil dari klasifikasi SVM memberikan nilai akurasi pada setiap nilai  $C$  kemudian dilakukan evaluasi model menggunakan *confusion matrix*. Dihitung juga nilai *precision*, *recall*, dan *F1-score*. *Precision* digunakan untuk menilai prediksi hasil positif benar dari semua hasil prediksi positif. *Recall* digunakan untuk menilai hasil prediksi positif benar dari semua hasil prediksi positif benar. Adapun *F1-score* adalah rata-rata harmoni dari *precision* dan *recall*.

Dalam artikel ini digunakan empat metode kernel dan empat nilai parameter *cost*. Metode kernel tersebut adalah *linear*, RBF, polinomial, dan sigmoid. Nilai parameter *cost* yang digunakan adalah 1, 10, 100, dan 1000. Digunakan *default* parameter untuk parameter *degree* pada kernel polinomial dan parameter *gamma* pada kernel *non-linear*.

**Tabel 4.** Daftar nilai akurasi

Kernel	Nilai $C$			
	1,0	10,0	100,0	1000,0
<i>Linear</i>	0,9522	1,0000	0,9932	0,9932
RBF	0,9021	0,9499	0,9613	0,9681
Polinomial	0,8861	0,8838	0,9043	0,9339
Sigmoid	0,7995	0,7995	0,8018	0,7995

Dari Tabel 4 didapatkan nilai akurasi tertinggi untuk masing-masing kernel. Kernel *linear* mendapatkan akurasi tertinggi pada nilai  $C$  10 dengan akurasi 100%. Kernel RBF akurasi tertinggi dicapai oleh nilai  $C$  1000 yaitu sebesar 96,81%. Kernel polinomial

memberikan hasil akurasi terbesar yaitu 93,33% pada nilai  $C$  1000. Kernel sigmoid meraih akurasi tertinggi pada nilai  $C$  100 dengan nilai akurasi 80,18%.

Secara keseluruhan, model bekerja dengan baik terutama pada kernel *linear* dan RBF yang mendapatkan hasil diatas 90% di semua nilai  $C$  meskipun kernel polinomial juga mendapat nilai akurasi yang baik. Kernel sigmoid mendapatkan hasil akurasi paling minimum dibanding kernel lainnya dengan akurasi dibawah 80% pada nilai  $C$  1, 10, dan 1000. Dapat juga dilihat bahwa nilai  $C$  100 memberikan nilai akurasi diatas 80% di seluruh kernel.

### 3.2 Evaluasi Model

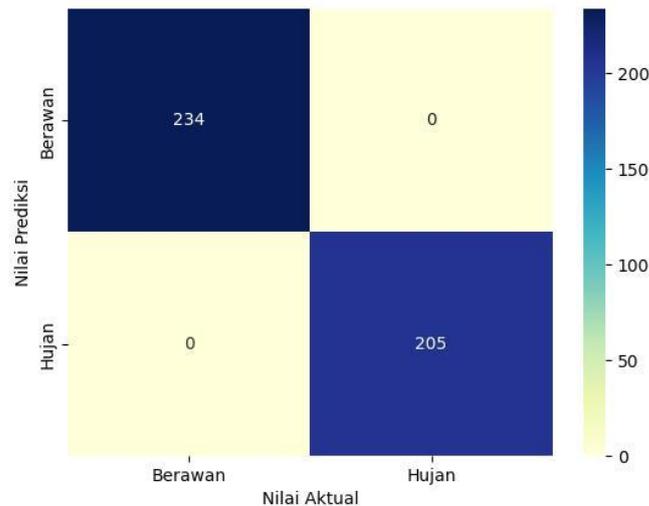
Tabel 5 memberikan hasil evaluasi menggunakan *confusion matrix* dengan *standard metrics precision, recall, dan F1-Score* pada masing-masing kernel dan nilai  $C$ . Evaluasi yang paling baik ditunjukkan oleh kernel *linear*, selanjutnya kernel RBF dan polinomial juga memberikan hasil yang bagus. Kernel sigmoid memberikan hasil evaluasi yang kurang bagus dibanding kernel lainnya dengan nilai evaluasi berkisar diantara 80%.

**Tabel 5.** Daftar evaluasi

Kernel	Nilai $C$	<i>Precision</i>		<i>Recall</i>		<i>F1-Score</i>	
		Berawan	Hujan	Berawan	Hujan	Berawan	Hujan
<i>Linear</i>	1,0	0,92	1,00	1,00	0,90	0,96	0,95
	10,0	1,00	1,00	1,00	1,00	1,00	1,00
	100,0	0,99	1,00	1,00	0,99	0,99	0,99
	1000,0	1,00	0,99	0,99	1,00	0,99	0,99
RBF	1,0	0,88	0,94	0,95	0,85	0,91	0,89
	10,0	0,93	0,98	0,98	0,91	0,95	0,94
	100,0	0,95	0,97	0,98	0,94	0,96	0,96
	1000,0	0,97	0,97	0,97	0,96	0,97	0,97
Polinomial	1,0	0,90	0,87	0,88	0,89	0,89	0,88
	10,0	0,90	0,86	0,88	0,89	0,89	0,88
	100,0	0,90	0,91	0,93	0,88	0,91	0,90
	1000,0	0,92	0,95	0,96	0,90	0,94	0,93
Sigmoid	1,0	0,83	0,77	0,79	0,81	0,81	0,79
	10,0	0,83	0,77	0,79	0,81	0,81	0,79
	100,0	0,83	0,77	0,79	0,82	0,81	0,79
	1000,0	0,83	0,77	0,79	0,82	0,81	0,79

Berdasarkan Tabel 4, akurasi tertinggi dicapai oleh kernel *linear* dengan nilai  $C$  10 sebesar 100%. Adapun plot *heatmap confusion matrix* untuk kernel *linear* dengan nilai  $C$  10 dapat dilihat di Gambar 2.

Model dengan kernel *linear* dan nilai  $C$  10 memberikan akurasi 100% pada data uji. Hal ini berarti skor *precision* 100% menunjukkan bahwa setiap prediksi benar adalah benar aktual sedangkan skor *recall* 100% menunjukkan bahwa setiap hasil benar telah diprediksi oleh model dengan benar atau tidak ada *false negative*. Selain itu, *F1-score* sebesar 100% menunjukkan bahwa kedua skor *precision* dan *recall* adalah sempurna. Hasil evaluasi ini menunjukkan bahwa tidak ada klasifikasi yang salah dan *hyperplane* yang memisahkan kelas berawan dan kelas hujan pada data uji sudah optimal.



**Gambar 2.** *Confusion Matrix* kernel linear,  $C = 10$

#### 4 KESIMPULAN

Berdasarkan hasil analisis model SVM pada klasifikasi cuaca Kabupaten Sleman dapat disimpulkan bahwa model SVM memberikan hasil yang bagus pada penggunaan kernel *linear* dan RBF dengan nilai akurasi di atas 90% pada seluruh nilai  $C$  dengan kernel *linear* sedikit lebih unggul dibanding kernel RBF. Kernel polinomial juga memberikan hasil yang baik yaitu akurasi di atas 90% pada nilai  $C$  sebesar 100 dan 1000. Akurasi kernel sigmoid pada nilai  $C$  100 memberikan nilai akurasi yang cukup yaitu sebesar 80,18% dengan nilai  $C$  lainnya berada di bawah 80%. Model klasifikasi prediksi cuaca paling akurat yang dapat digunakan untuk kasus klasifikasi cuaca Kabupaten Sleman adalah algoritma SVM dengan parameter kernel *linear* dengan nilai  $C$  10. Dengan akurasi 100%, menunjukkan bahwa model optimal dalam memberikan prediksi klasifikasi cuaca berdasarkan curah hujan pada kasus tersebut.

#### DAFTAR PUSTAKA

- Awad, M. & Khanna, R. (2015). Support Vector Machines for Classification. Dalam *Efficient Learning Machines*. Apress, Berkeley, CA. doi:10.1007/978-1-4302-5990-9\_3
- Belete, D. M., & Huchaiyah, M. D. (2021). Grid Search in Hyperparameter Optimization of Machine Learning Models for Prediction of HIV/AIDS Test Result. *INTERNATIONAL JOURNAL OF COMPUTERS AND APPLICATIONS*. doi:10.1080/1206212X.2021.1974663
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer New York, NY.
- BMKG. (2024). *Informasi Cuaca - Probabilistik Curah Hujan 20 mm (tiap 24 jam)*. Diambil kembali dari Badan Meteorologi Klimatologi dan Geofisika: <https://www.bmkg.go.id/>
- BMKG. (2024). *Prediksi Musim Kemarau 2024 di Indonesia*. Jakarta: Badan Meteorologi Klimatologi dan Geofisika.
- Dasmasela, R., Tomasouw, B. P., & Leleury, Z. A. (2022). Penerapan Metode Support Vector Machine (SVM). *PARAMETER - JURNAL MATEMATIKA, STATISTIKA DAN TERAPANNYA*. doi:doi.org/10.30598/parameterv1i2pp111-122
- Fremmuzar, P. & Baita, A. (2023). Uji Kernel SVM dalam Analisis Sentimen Terhadap Layanan Telkomsel di Media Sosial Twitter. *KOMPUTIKA*. doi:10.34010/komputika.v12i2.9460
- James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2013). *An Introduction to Statistical Learning: with Application in R*. Springer New York, NY.

- Jayadianti. (2020). Metode Komparasi Artificial Neural Network Pada Prediksi Curah Hujan. *Jurnal Tekno Insentif*. doi:10.36787/jti.v14i2.150
- Lomboy, K. E. & Hernandez, R. M. (2021). A comparative performance of breast cancer classification using hyper-parameterized machine learning models. *International Journal of Advanced Technology and Engineering Exploration*. doi:10.19101/IJATEE.2021.874380
- Mardiansyah, R. Y., Kurniawan, B., Soekirno, S., Nuryanto, D. E., & Satria, H. (2022). Artificial Intelligent For Rainfall Estimation In. *IOP Conference Series: Earth and*. doi:10.1088/1755-1315/1105/1/012024
- Pratiwi, N. & Setyawan, Y. (2021). Analisis Akurasi dari Perbedaan Fungsi Kernel dan Cost pada Support Vector Machine Studi Kasus Klasifikasi Curah Hujan di Jakarta. doi:10.14710/jfma.v4i2.11691
- Pridiptama, R. P., Wasono, W., & Amijaya, F. D. (2024). Perbandingan Algoritma Support Vector Machine dan Naive Bayes pada Klasifikasi Penyakit Tekanan Darah Tinggi. *Basis*. doi:10.30872/basis.v3i1.1264
- Putra, A. P., Debatara, N. N., & Kusnandar, D. (2020). Tingkat Akurasi Klasifikasi Jarak Kelahiran di Kampung Keluarga Berencana (KB) dengan Metode SVM. *BIMASTER*. doi:10.26418/bbimst.v9i3.40891
- Sunarmi, N., Kumailia, E. N., Nurfaiza, N., Nikmah, A. K., Aisyah, H. N., Sriwahyuni, I., & Laily, S. N. (2022). Analisis Faktor Unsur Cuaca terhadap Perubahan Iklim Di Kabupaten Pasuruan. *Newton-Maxwell Journal of Physics*. doi:10.33369/nmj.v3i2.23380.