

## Analisis Peramalan Curah Hujan Menggunakan Metode Extreme Learning Machine (Studi Kasus: Stasiun Klimatologi Jawa Timur)

Anjas Pramono<sup>1</sup>, Arief Andy Soebroto<sup>2</sup>

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya  
Email: <sup>1</sup>anjas\_pramono@student.ub.ac.id, <sup>2</sup>ariefas@ub.ac.id

### Abstrak

Penelitian ini fokus pada analisis peramalan curah hujan di Indonesia, sebuah negara tropis yang terdapat curah hujan tinggi sehingga menyebabkan banjir dan tanah longsor. Stasiun Klimatologi Jawa Timur, terletak di Kabupaten Malang, merupakan pusat informasi cuaca dan sering terkena dampak banjir. Studi ini mengidentifikasi pola data historis secara sistematis untuk peramalan yang akurat. Metode peramalan yang digunakan adalah Jaringan Saraf Tiruan, khususnya Extreme Learning Machine (ELM). ELM dianggap efektif dengan tingkat kesalahan rendah dan kecepatan pelatihan yang tinggi. Penelitian ini membandingkan kinerja ELM dibandingkan metode tradisional seperti KNN dan SVM, menunjukkan superioritas ELM dalam kecepatan dan kinerja komputasi. Sebuah studi kasus menggunakan judul "Analisis Peramalan Curah Hujan Menggunakan Metode Extreme Learning Machine (Studi Kasus: Stasiun Klimatologi Jawa Timur)" menyoroti penerapan ELM dalam konteks ini. Temuan penelitian ini dapat mendukung upaya pencegahan banjir melalui peramalan yang lebih akurat. Kinerja peramalan curah hujan dengan metode ELM memperoleh nilai Mean Squared Error (MSE) 0,021 rasio parameter data pelatihan dan data pengujian sebesar 50% - 50% dan jumlah hidden neuron sebanyak 10 neuron.

**Kata kunci:** analisis peramalan, curah hujan, jaringan syaraf tiruan, ELM, MSE

### Abstract

*This research focuses on analyzing rainfall forecasting in Indonesia, a tropical country with high rainfall that causes floods and landslides. The East Java Climatology Station, located in Malang Regency, is a weather information center and is often affected by flooding. This study identifies historical data patterns systematically for accurate forecasting. The forecasting method used is Artificial Neural Network, specifically Extreme Learning Machine (ELM). ELM is considered effective with low error rate and high training speed. This research compares the performance of ELM against traditional methods such as KNN and SVM, showing the superiority of ELM in computational speed and performance. A case study using the title "Rainfall Forecasting Analysis Using Extreme Learning Machine Method (Case Study: East Java Climatology Station)" highlights the application of ELM in this context. The findings of this study can support flood prevention efforts through more accurate forecasting. The performance of rainfall forecasting with the ELM method obtained a Mean Squared Error (MSE) value of 0.021, a ratio of training data parameters and test data of 50% - 50% and the number of hidden neurons of 10 neurons.*

**Keywords:** forecast, rainfall, artificial neural network, ELM, MSE

## 1. PENDAHULUAN

Indonesia adalah negara tropis yang menerima banyak curah hujan (Sidiq, 2018). Bencana alam seperti banjir dan tanah longsor terjadi karena Indonesia memiliki curah hujan tinggi. Stasiun Klimatologi Jawa Timur yang berada di Kecamatan Karangploso, Kabupaten Malang berfungsi sebagai pusat informasi data cuaca dimana salah satu lokasi yang sering terkena dampak banjir.

Identifikasi pola data secara sistematis

terhadap data historis adalah subjek dari studi peramalan (Wahyuni, et al., 2016). Di era sekarang, peramalan membantu menentukan proyeksi curah hujan secara akurat berdasarkan tren data yang sistematis. Peningkatan upaya pencegahan banjir dapat dicapai dengan peramalan yang tepat.

Peramalan dilakukan menggunakan teknik jaringan saraf tiruan. Teknik pemodelan data yang kuat yang mampu menggambarkan interaksi input- output yang rumit adalah jaringan saraf tiruan (Sudarsono, 2016). Karena

Jaringan Syaraf Tiruan dapat menangani model yang rumit, mereka membantu dalam peramalan yang tepat.

Extreme Learning Machine (ELM) memiliki kesalahan yang rendah dan keunggulan dalam kecepatan pelatihan, adalah salah satu teknik Jaringan Syaraf Tiruan (Simamora, et al., 2019). ELM bekerja lebih efektif daripada pembelajaran mesin tradisional (Wang, et al., 2021).

Menurut studi oleh Chy dan Rahaman (2019), ELM berkinerja lebih baik daripada KNN dan SVM. Selain itu, penelitian oleh Ahmad et al. (2018) menunjukkan bahwa ELM mengungguli teknik pembelajaran mesin konvensional seperti SVM dan Random Forest dalam hal kecepatan dan kinerja komputasi. "Analisis Peramalan Curah Hujan Menggunakan Metode Extreme Learning Machine (Studi Kasus: Stasiun Klimatologi Jawa Timur)" adalah penelitian yang dapat dilakukan dengan metode ELM.

## 2. LANDASAN KEPUSTAKAAN

### 2.1. Tinjauan Pustaka

(Chy & Rahaman, 2019) fokus pada topik perbandingan kinerja metode KNN, SVM, dan Extreme Learning Machine (ELM). Pengujian dilakukan dengan menggunakan dataset citra penyakit Sickle Cell Anemia (SCA) dengan hasil metode KNN meraih precision 90%, accuracy 73,33%, f-score 81,81%, dan waktu komputasi 10,15 detik. Pengujian dengan metode SVM meraih precision 95,45%, accuracy 83,33%, f-score 88,89%, dan waktu komputasi 5,91 detik. Pengujian dengan metode ELM meraih precision 95,45%, accuracy 87,73%, f-score 91,3%, dan waktu komputasi 0,05 detik.

(Ahmad, et al., 2018) fokus pada topik perbandingan kinerja metode SVM (linear), SVM (RBF), Random Forest, ELM dengan menggunakan dataset deteksi gangguan keamanan cybersecurity dengan hasil metode SVM (linear) meraih accuracy 98,8%, precision 97,5%, dan recall 97,8%. Pengujian dengan metode SVM (RBF) meraih accuracy 98,2%, precision 97,3%, dan recall 97,4%. Pengujian dengan metode Random Forest meraih accuracy 97,7%, precision 97,9%, dan recall 96,4%. Pengujian dengan metode ELM mendapatkan hasil accuracy 99,5%, precision 98,6%, dan recall 98,6%.

(Zainuddin, et al., 2019) fokus pada topik

kinerja metode ELM dan KNN dengan menggunakan dataset sinyal EEG pada anak-anak penderita disleksia dengan hasil metode KNN meraih accuracy 83% pada dataset db2 dan db4. Pengujian dengan metode ELM meraih accuracy 89% pada dataset db2, dan accuracy 83% pada dataset db4.

### 2.2. Curah Hujan

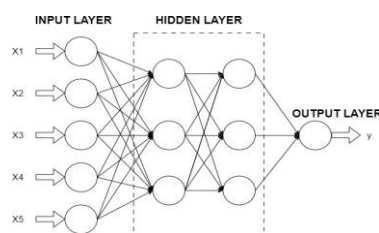
Curah hujan didefinisikan sebagai air yang ditampung di suatu area datar tanpa mengalir, meresap, atau menguap (Simamora, et al., 2019). Ketinggian curah hujan dinyatakan dalam milimeter.

### 2.3. Peramalan

Peramalan adalah statistik umum yang membantu dalam membuat keputusan tentang penjadwalan produksi, transportasi, dan perencanaan strategis jangka panjang. Karena peramalan sangat penting dalam banyak hal, yang menjadi bagian utama waktu pengambilan keputusan manajemen. Berdasarkan tujuan khususnya organisasi, bisnis modern memerlukan prediksi jangka pendek, menengah, dan panjang (Hyndman & Athanasopoulos, 2018).

### 2.4. Jaringan Syaraf Tiruan

Jaringan Syaraf Tiruan adalah metode yang tergolong didalam jenis Kecerdasan Buatan. Jaringan Syaraf Tiruan mengolah informasi melalui neuron yang setiap sambungan antar neuron memiliki bobot (Kayabaşı, et al., 2018). Jaringan Saraf Tiruan dapat mengumpulkan informasi tentang sampel, membuat generalisasi, dan kemudian memutuskan pada sampel tersebut menggunakan informasi yang dipelajari (Kayabaşı, et al., 2018). Jaringan Syaraf Tiruan memiliki tiga lapisan secara umum, yaitu Neuron Masukkan, Neuron Tersembunyi, dan Neuron Keluaran (Kayabaşı, et al., 2018). Arsitektur Jaringan Syaraf Tiruan ditampilkan pada Gambar 1.



Gambar 1 Jaringan Syaraf Tiruan

### 2.5. Preprocessing

Preprocessing adalah tahap proses pengolahan fitur yang relevan untuk proses peramalan (Zhao, et al., 2006). Preprocessing mendukung untuk meningkatkan performa peramalan.

Missing values adalah kejadian data yang disebabkan karena faktor manusia, masalah teknis, dan sebagainya (Petrusevich, 2021). Penanganan data pada missing values dieksekusi bertujuan untuk mendapatkan nilai rata-rata dari nilai indeks data setelahnya dan nilai indeks data sebelumnya (Petrusevich, 2021).

Normalisasi adalah strategi preprocessing yang memiliki implikasi signifikan terhadap accuracy peramalan (Panigrahi, et al., 2013). Normalisasi Min-Max adalah salah satu jenis normalisasi. Tujuan dari normalisasi min-max adalah untuk memberikan perbandingan nilai yang lebih seimbang antara data sebelum dan setelah proses dengan mentransformasikan data asli secara linear (Henderi, et al., 2021). Proses penghitungan normalisasi ditunjukkan pada rumus di Persamaan 1.

$$X = \frac{(X - X_{min})}{(X_{max} - X_{min})} \quad (1)$$

### 2.6. Denormalisasi

Denormalisasi adalah proses  $i=1$  restruktur keluaran Jaringan Syaraf Tiruan menjadi keluaran data asli (Herawati & Latif, 2020). Proses penghitungan denormalisasi ditunjukkan pada rumus di Persamaan 2.

$$X = (X'(X_{max} - X_{min})) + X_{min} \quad (2)$$

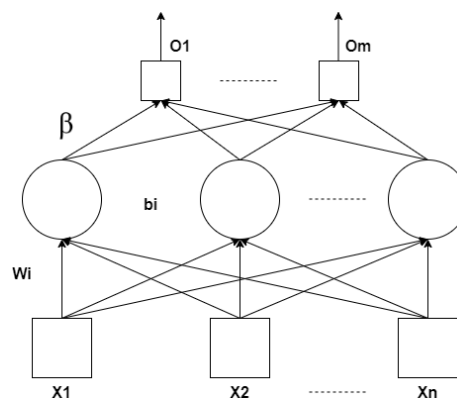
### 2.7. Fungsi Aktivasi

Fungsi aktivasi adalah metode yang dipakai pada Jaringan Syaraf Tiruan untuk mengubah sinyal input menjadi sinyal output (Sharma, et al., 2020). Fungsi aktivasi dibutuhkan oleh Jaringan Syaraf Tiruan karena keluaran di Jaringan Syaraf Tiruan hanya berbentuk fungsi linier sederhana yang memiliki keterbatasan dalam mengenali kompleksitas data (Sharma, et al., 2020).

### 2.8. Extreme Learning Machine (ELM)

ELM adalah jenis Jaringan Syaraf Tiruan yang memiliki arsitektur Single- Hidden Layer Feedforward Neural Networks (SLFNs) yang hanya memiliki lapisan tersembunyi tunggal

dengan mendapatkan masukan bobot dan bias secara random (Huang, et al., 2006). Arsitektur ditunjukkan pada Gambar 2.



Gambar 2 Extreme Learning Machine

Proses training bertujuan untuk melakukan proses pembelajaran terhadap dataset. Tahapan proses pelatihan pada Extreme Learning Machine dilakukan dengan Langkah pertama parameter bobot dan bias harus terlebih dahulu dimasukan dengan acak menggunakan rentang nilai dari -1 hingga 1. Langkah kedua adalah melakukan perhitungan masukan lapisan tersembunyi. Pada tahapan ini proses menghitung keluaran di lapisan tersembunyi ditampilkan pada Persamaan 3.

$$H_{init\ ij} = (\sum_{j=1}^n x_j w_j^T) + b \quad (3)$$

Langkah ketiga adalah menghitung matriks keluaran lapisan tersembunyi pada Persamaan 4.

$$H(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

Langkah keempat adalah menghitung Matriks Moore-Penrose Generalized Inverse pada Persamaan 5.

$$H^+ = (H(x)^T \cdot H(x))^{-1} \cdot H(x)^T \quad (5)$$

Langkah kelima adalah menghitung keluaran bobot pada Persamaan 6.

$$\beta = H^+ \cdot T \quad (6)$$

Langkah keenam adalah menghitung hasil akhir peramalan. Rumus perhitungan ditunjukkan pada Persamaan 7.

$$O = H(x) \cdot \beta \quad (7)$$

Proses testing adalah prosedur pengujian menggunakan pendekatan Extreme Learning Machine (ELM) untuk menilai hasil dari proses pelatihan. Tahapan proses testing seperti yang dijelaskan oleh (Huang, et al., 2006) adalah

masukan bobot dan bias dari tahapan proses training menjadi dasar masukan bobot dan bias pada tahapan testing. Tahapan kedua penghitungan hidden layer sesuai pada Persamaan 3 dan Persamaan 4. Hasil akhir peramalan menggunakan Persamaan 7 dengan nilai keluaran hidden layer pada nomor 2 dan matriks keluaran bobot dengan nilai yang diambil dari tahapan proses training.

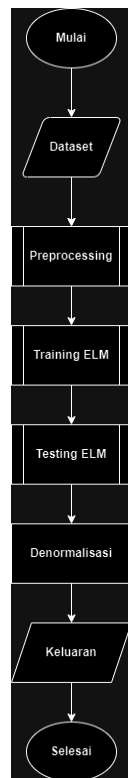
**2.9. Mean Squared Error (MSE)**

Salah satu teknik untuk mengetahui nilai kesalahan adalah dengan memanfaatkan Mean Squared Error (MSE) (Widiarti, et al., 2017). Prosedur peramalan akan semakin akurat jika nilai MSE semakin rendah. Rumus untuk mengetahui nilai kesalahan menggunakan MSE adalah Persamaan 8.

$$MSE = \sum_{i=1}^n \frac{(y_i - t_i)^2}{n} \tag{8}$$

**3. METODOLOGI PENELITIAN**

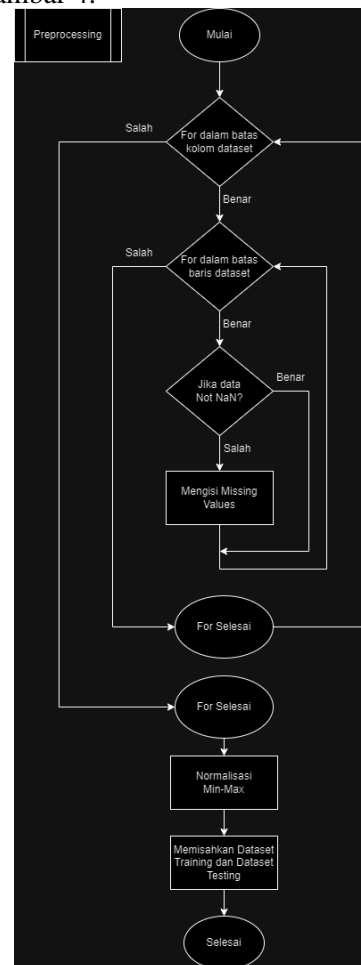
Metodologi penelitian yang dilakukan dengan diagram alur penelitian agar tahapan yang dilakukan terlaksana dengan baik. Metodologi penelitian ini disampaikan dengan alur penelitian yang akan disajikan pada Gambar 3.



Gambar 3 Diagram Alir Penelitian

Penelitian ini dilaksanakan dengan dataset curah hujan Stasiun Klimatologi Jawa Timur dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG). Selanjutnya dilakukan tahapan preprocessing. Tahapan preprocessing melakukan pengolahan terhadap data missing values, selanjutnya melakukan normalisasi, denormalisasi, dan memisahkan data menjadi data latih dan data uji. Tahap selanjutnya melakukan tahapan training dengan menggunakan metode Extreme Learning Machine (ELM). Setelah proses training dilakukan, selanjutnya dilakukan proses testing ELM dengan menggunakan bobot neuron hasil training ELM. Denormalisasi digunakan untuk mengembalikan nilai hasil peramalan ke bentuk asli. Keluaran adalah hasil akhir peramalan yang telah diubah ke bentuk asli.

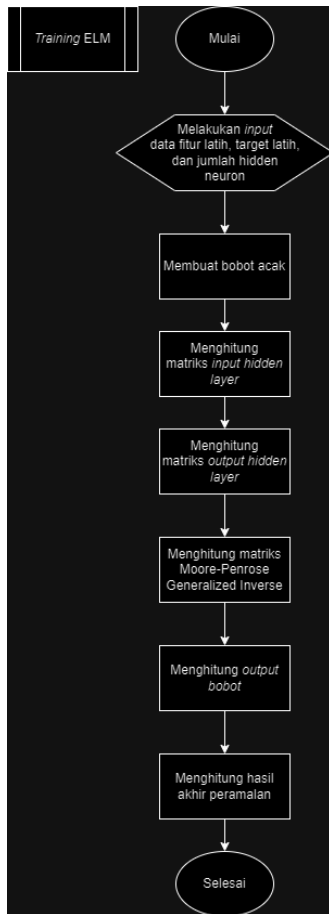
Preprocessing dilakukan untuk mempersiapkan data yang akan diproses pada tahap pelatihan Extreme Learning Machine. Diagram alir untuk preprocessing ditunjukkan pada Gambar 4.



Gambar 4 Diagram Alir Preprocessing

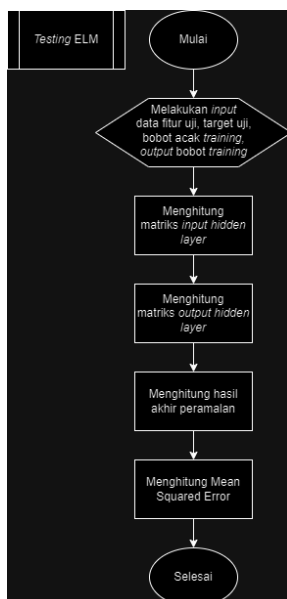
Proses untuk melatih data curah hujan dengan pendekatan ELM, prosedur pelatihan

ELM dijalankan. Diagram alir pelatihan ELM ditampilkan pada Gambar 5.



Gambar 5 Diagram Alir Training ELM

Prosedur pengujian ELM digunakan untuk mengevaluasi data curah hujan dengan menggunakan pendekatan ELM. Diagram alir pengujian ELM ditampilkan pada Gambar 6.



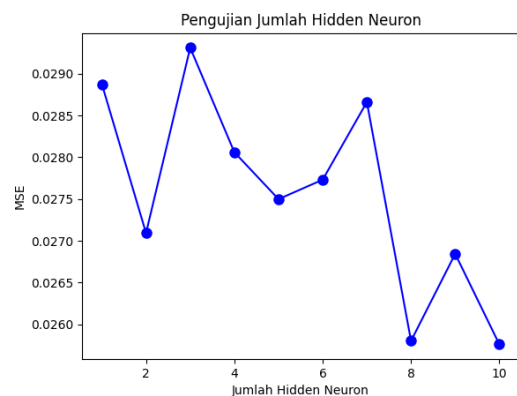
Gambar 6 Diagram Alir Testing ELM

#### 4. HASIL DAN PEMBAHASAN

Tahapan pengujian jumlah lapisan tersembunyi untuk menemukan jumlah neuron tersembunyi yang lebih baik dengan hasil Mean Squared Error (MSE) yang terbaik adalah tujuan dari tahap pengujian jumlah neuron tersembunyi. Data hasil dari proses pengujian jumlah neuron tersembunyi ditampilkan pada Tabel 1 dan Gambar 7.

Tabel 1 Pengujian Jumlah Lapisan Tersembunyi

Jumlah Hidden Neuron	MSE
1	0,029
2	0,028
3	0,029
4	0,028
5	0,027
6	0,028
7	0,029
8	0,026
9	0,027
10	0,026



Gambar 7 Grafik Pengujian Jumlah Hidden Neuron

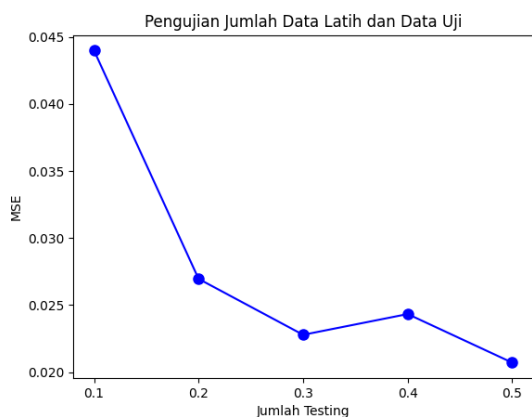
Berdasarkan Tabel 1 dan Gambar 7 hasil ketika dilakukan pengujian jumlah lapisan tersembunyi yang terlihat bahwa nilai MSE terbaik didapatkan ketika jumlah lapisan tersembunyi berjumlah 10 dengan nilai MSE 0,026. Hasil pengujian jumlah lapisan tersembunyi dari neuron sebanyak 1 hingga 10 terlihat bahwa ketika menggunakan jumlah hidden neuron lebih sedikit menyebabkan underfitting. Underfitting terjadi karena proses pelatihan gagal mengenali pola secara baik yang disebabkan karena proses pelatihan kurang kompleks.



Tahapan pengujian pada jumlah data pelatihan dan data pengujian untuk mendapatkan komposisi data pelatihan dan data pengujian yang optimal dengan hasil Mean Squared Error (MSE) yang terbaik. Hasil pengujian jumlah data uji dan data latih ditampilkan pada Tabel 2 dan Gambar 8.

Tabel 2 Pengujian Jumlah Data Latih dan Data Uji

Persentase Data Latih	Persentase Data Uji	MSE
50%	50%	0,021
60%	40%	0,024
70%	30%	0,023
80%	20%	0,027
90%	10%	0,044



Gambar 8 Grafik Pengujian Jumlah Data Latih dan Data Uji

Menurut Tabel 2 dan Gambar 8 hasil dari pengujian jumlah data latih dan data uji terlihat bahwa nilai MSE terbaik didapatkan ketika persentase data dengan nilai MSE 0,021. Hasil pengujian jumlah data pelatihan dan data pengujian terlihat bahwa ketika menggunakan persentase testing yang lebih besar menghasilkan nilai MSE yang lebih baik disebabkan karena overfitting. Overfitting terjadi karena proses training terlalu kompleks diakibatkan data pelatihan yang terlalu besar dan menghasilkan peramalan secara baik saat proses training tetapi memiliki kinerja yang buruk ketika mengenali data baru di proses testing.

**5. PENUTUP**

Kesimpulan yang diraih setelah tahapan proses pengujian dan pembahasan mengenai "Analisis Peramalan Curah Hujan Menggunakan Metode Extreme Learning Machine (Studi Kasus: Stasiun Klimatologi Jawa Timur)":

"Analisis Peramalan Curah Hujan Menggunakan Metode Extreme Learning Machine (Studi Kasus: Stasiun Klimatologi Jawa Timur)", yang diawali dengan preprocessing, dilanjutkan dengan pelatihan dan pengujian ELM, dan diakhiri dengan denormalisasi untuk mendapatkan hasil peramalan curah hujan

Penelitian "Analisis Peramalan Curah Hujan Menggunakan Metode Extreme Learning Machine (Studi Kasus: Stasiun Klimatologi Jawa Timur)" menghasilkan hasil terbaik dengan 10 hidden neuron dan rasio data pelatihan dan pengujian sebesar 50%-50% dengan 0,021 merupakan hasil perhitungan nilai Mean Squared Error (MSE).

Penelitian untuk mendapatkan hasil peramalan yang cepat dan akurat dengan total data yang lebih besar secara ukuran dengan jangka waktu yang lebih panjang, maka penelitian yang berjudul "Analisis Peramalan Curah Hujan Menggunakan Metode Extreme Learning Machine (Studi Kasus: Stasiun Klimatologi Jawa Timur)" ini dapat ditransformasikan menjadi peramalan yang dilakukan setiap hari secara real time dengan menggunakan teknik deep learning.

**6. DAFTAR REFERENSI**

Aggarwal, C. C., 2018. *Neural Networks and Deep Learning A Textbook*. New York: Springer.

Ahmad, I., Basher, M., Iqbal, M. J. & Rahim, A., 2018. Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection. *SPECIAL SECTION ON SURVIVABILITY STRATEGIES FOR EMERGING WIRELESS NETWORKS*, Volume 6, pp. 33789 - 33795.

Chy, T. S. & Rahaman, M. A., 2019. A Comparative Analysis by KNN, SVM & ELM Classification to Detect Sickle Cell Anemia. Dhaka, 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST).

Henderi, Wahyuningsih, T. & Rahwanto, E., 2021. Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (KNN) Algorithm to Test the Accuracy of Types of Breast Cancer. *International*

- Journal of Informatics and Information System, 4(1), pp. 13-20.
- Herawati, S. & Latif, M., 2020. Forecasting tourist visits using data decomposition technique and learning optimization of artificial neural network. Surabaya, Journal of Physics: Conference Series.
- Huang, G.-B., Zhu, Q.-Y. & Siew, C.-K., 2006. Extreme learning machine: Theory and applications. Neurocomputing, 70(1-3), pp. 489-501.
- Hyndman, R. J. & Athanasopoulos, G., 2018. Forecasting: Principles and Practice. 2nd ed. Melbourne: Otexts.
- Kayabaşı, A., Yıldız, B. & Aslan, M. F., 2018. Comparison of ELM and ANN on EMG Signals Obtained for Control of Robotic-Hand. Iasi, ECAI 2018 - International Conference – 10th Edition.
- Khair, U., Fahmi, H., Hakim, S. A. & Rahim, R., 2017. Forecasting Error Calculation with Mean Absolute Deviation and Mean Absolute Percentage Error. Melaka, International Conference on Information and Communication Technology (IconICT).
- Panigrahi, S., Karali, Y. & Behera, H. S., 2013. Normalize Time Series and Forecast using Evolutionary Neural Network. International Journal of Engineering Research & Technology (IJERT), 2(9).
- Petrusevich, D. A., 2021. Review of missing values procession methods in time series data. Journal of Physics: Conference Series, 1889(3), p. 032009.
- Sharma, S., Sharma, S. & Athaiya, A., 2020. ACTIVATION FUNCTIONS IN NEURAL NETWORKS. International Journal of Engineering Applied Sciences and Technology, 4(12), pp. 310-316.
- Sidiq, M., 2018. Forecasting Rainfall with Time Series Model. Bandung, IOP Conference Series.
- Simamora, R. J. D., Tibyani & Sutrisno, 2019. Peramalan Curah Hujan Menggunakan Metode Extreme Learning Machine. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 3(10), pp. 9670-9676.
- Sudarsono, A., 2016. JARINGAN SYARAF TIRUAN UNTUK MEMPREDIKSI LAJU PERTUMBUHAN PENDUDUK MENGGUNAKAN METODE BACPROPAGATION (STUDI KASUS DI KOTA BENGKULU). Jurnal Media Infotama, 12(1).
- Wahyuni, N. P. M. S., Sumarjaya, I. W. & Srinadi, I. G. A. M., 2016. PERAMALAN CURAH HUJAN MENGGUNAKAN METODE ANALISIS SPEKTRAL. E-Jurnal Matematika, 5(4), pp. 183-193.
- Wang, J., Lu, S., Wang, S.-H. & Zhang, Y.-D., 2021. A review on extreme learning machine. Multimedia-based Healthcare Systems using Computational Intelligence, Volume 81, pp. 41611-41660.
- Widiarti, Pertiwi, R. R. & Sutrisno, A., 2017. Perbandingan Mean Squared Error (MSE) Metode Prasad-Rao dan Jiang-Lahiri-Wan Pada Pendugaan Area Kecil. Jakarta, Seminar Nasional Teknoka.
- Zainuddin, A. Z., Mansor, W., Lee, K. Y. & Mahmoodin, Z., 2019. Comparison of Extreme Learning Machine and K-Nearest Neighbour Performance in Classifying EEG Signal of Normal, Poor and Capable Dyslexic Children. Berlin, 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).
- Zhao, J. H., Dong, Z. & Xu, Z., 2006. Effective Feature Preprocessing for Time Series Forecasting. Xi'an, International Conference on Advanced Data Mining and Applications.