IDENTIFIKASI PROFIL KONDISI PENGERINGAN BERBASIS SENSOR SUHU DAN KELEMBAPAN MENGGUNAKAN K-MEANS CLUSTERING

Musthofa Galih Pradana[1)], Nurhuda Maulana[1)], I Wayan Rangga Pinastawa[1)], Fadhli Suko Wiryanto[2)] , Nurul Afifah Arifuddin[1)], Noor Falih[1)], Retno Dwi Nyamiati[3)], Ryan Setya Budi[4)]

[1]Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta
[2]Fakultas Ekonomi dan Bisnis, Universitas Pembangunan Nasional Veteran Jakarta
[3]Fakultas Teknik Industri, Universitas Pembangunan Nasional Veteran Yogyakarta
[4]Fakultas Ekonomi dan Bisnis, Universitas Pembangunan Nasional Veteran Yogyakarta

E-mail: musthofagalihpradana@upnvj.ac.id

Abstract

Drying process is one of the important stages in product handling that requires control of environmental conditions, especially temperature and humidity. Changes in these two parameters during the drying process can form different condition patterns and can potentially be used to identify process characteristics. This study aims to analyze temperature and humidity patterns during the drying process using the clustering approach without labels. The data used consisted of 30 trials of the drying process with 1,004 observations, which recorded temperature and humidity periodically. The research stages include pre-processing of data, extraction of statistical features and characteristics of sensor changes, data normalization, and grouping using K-Means. The determination of the number of clusters was evaluated using the Elbow Method and Silhouette Score in candidates K=2 to K=6. The results of the evaluation showed that K=2 gave the highest Silhouette Score value of 0.2852, so it was used as the optimal number of clusters. The cluster profiles showed differences in characteristics mainly in the average temperature and humidity as well as the pattern of humidity changes during the drying process. Cluster 1 had an average temperature of 34.673 °C and a humidity of 58.101%, while Cluster 2 had an average temperature of 35.731 °C and a humidity of 60.163%. The results suggest that clustering can be used as an exploratory approach to identify patterns of environmental conditions in the drying process based on sensor data. These findings can be the basis for the analysis and development of data-driven drying process monitoring systems in future studies.

Keywords: *Clustering, K-Means, Drying Temperature, Humidity, Sensor.*

PENDAHULUAN

Tahapan pengeringan merupakan proses yang penting dalam alur dan pengolahan bahan berbasis serat alami untuk diolah ke dalam proses pembuatan kerajinan tangan. Kondisi lingkungan selama pengeringan adalah kunci utama dalam proses ini dimana hali ini dapat memengaruhi karakteristik dan kualitas produk yang dihasilkan. Suhu dan kelembapan relatif merupakan dua parameter lingkungan yang berperan dalam perpindahan air selama proses pengeringan. Perubahan kedua parameter tersebut berlangsung secara dinamis sehingga pengamatan menggunakan sensor dapat menghasilkan data yang lebih representatif terhadap kondisi proses dibandingkan

pengukuran secara periodik. Teknologi yang dapat membantu proses pada tahapan ini adalah dengan memanfaatan Internet of Things (IoT) (Pinheiro et al., 2023) yang memungkinkan parameter lingkungan selama proses pengeringan direkam secara kontinu. Data sensor yang dihasilkan tidak hanya dapat digunakan untuk pemantauan, tetapi juga dapat dianalisis untuk menemukan pola kondisi proses. Salah satu pendekatan yang dapat digunakan adalah clustering, yaitu teknik unsupervised learning yang mengelompokkan data berdasarkan kemiripan karakteristik tanpa memerlukan label kelas sebelumnya. K-Means merupakan salah satu metode clustering yang sederhana dan banyak digunakan karena proses komputasinya relatif ringan serta hasil kelompoknya mudah diinterpretasikan.

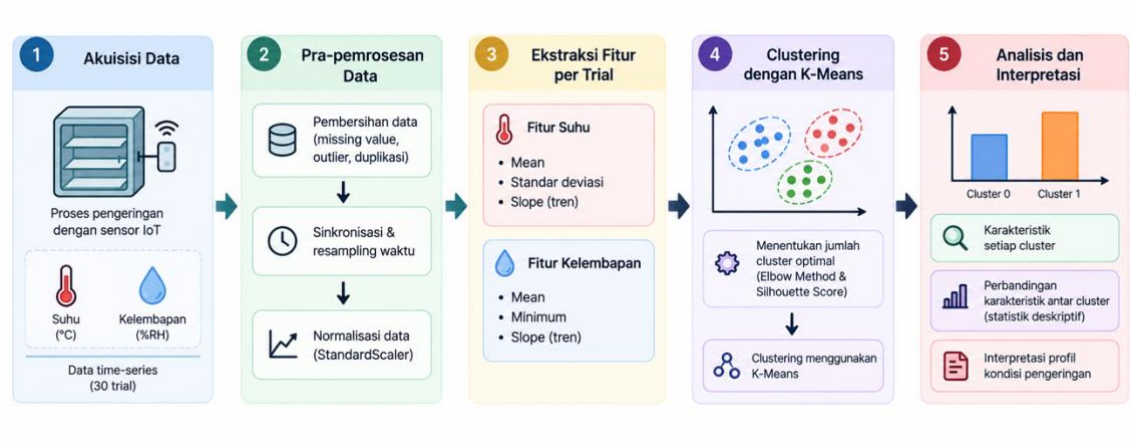
Beberapa penelitian terdahulu menunjukan bahwa proses pengolahan data sensor spesifiknya adalah sensor suhu dan kelembapan seperti penelitian pada pengeringan realtime (Jasmine & Triawati, 2022) dengan pemanfaatan suhu dan relative humidity dalam identifikasi lokasi yang ideal pada lingkungan pertanian (Dooyum Uyeh et al., 2022a). Hal ini mengindikasikan data sensor lingkungan mampu memiliki potensi proses analisis lebih lanjut. Analisis mendalam tersebut dapat dilakukan dengan pendekatan machine learning, dan tidak hanya terpaku pada hasil data sensor IoT (Novando & P, 2025) yang biasanya digunakan untuk monitoring dengan data suhu, kelembapan (Waworundeng, 2024) dimana data juga data analisis karakteristik dan pola (Utami & Hadi, 2025). Riset relevan lain, arah machine learning digunakan pada tahapan untuk melakukan prediksi kadar ait, optimasi proses maupun identifikasi karakteristik (Xing et al., 2024). Pengolahan data bisa mennggunakan pendekatan supervised (Galih Pradana et al., 2023) (Kumar et al., 2024) maupun unsupervised seperti K-Means (Rami & Malek, 2024) (Gudivaka et al., 2024)dalam menemukan kelompok/pola pada beberapa tahap proses pengeringan (Pal & Gope, 2024). Hal ini relevan dengan jenis data yang melakukan pengolahan data sensor, dengan alur tahapan preprocessing, feature selection, dan analisis (Liu & Qiu, 2026) (W. Hamaali et al., 2023). Penelitian relevan yang membandingkan berbagai metode clustering, termasuk K-Means, pada data sensor lingkungan seperti temperature dan humidity, dimana penelitian ini menunjukkan bahwa transformasi berbasis turunan/derivative dapat membantu menangkap dinamika temporal dalam data sensor (Ngo & Phan, 2025).

.

METODE PENELITIAN

2.1. Tahapan Penelitian

Penelitian yang dilakukan melalui beberapa tahapan dan langkah diantaranya : (1) Akuisisi Data dari proses pengeringan dengan sensor IoT dengan jumlah trial sebanyak 30, (2) Pra Pemrosesan Data dengan tahapan pembersihan data, sinkronisasi dan normalisasi data, (3) Ekstrasi Fitur dengan detail fitur suhu dan fitur kelembapan, (4) Clustering dengan K-Means dengan penentuan jumlah cluster optimal dan (5) Analisis Interpretasi dengan identifikasi karakteristik cluster, dan interpretasi profil. Adapun detail tahapan ditunjukan pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.2. Akuisisi Data

Data yang digunakan merupakan data hasil pemantauan proses pengeringan menggunakan sensor IoT. Dataset terdiri atas 30 trial, yaitu T01 sampai dengan T30, dengan jumlah observasi antara 31–36 observasi pada setiap trial. Setiap observasi merepresentasikan kondisi proses pada waktu tertentu. Parameter utama yang digunakan dalam penelitian adalah suhu (temperature_C) dan kelembapan relatif (humidity_RH_pct).

2.3. Pra Pemrosesan Data

Tahap pra-pemrosesan dilakukan untuk memastikan data sensor yang digunakan dalam analisis memiliki kualitas yang baik. Data yang digunakan merupakan data hasil pemantauan proses pengeringan yang terdiri atas variabel suhu (temperature) dan

kelembapan relatif (relative humidity). Pada tahap ini dilakukan pemeriksaan terhadap struktur data, nilai hilang (missing value), data duplikat, serta konsistensi tipe data. Hasil pemeriksaan menunjukkan bahwa data terdiri atas 30 trial dengan total 1.004 observasi, tanpa ditemukan data duplikat maupun nilai hilang. Selanjutnya, data dikelompokkan berdasarkan trial_id sebagai satuan pengamatan proses pengeringan.

2.4. Ekstrasi Fitur

Ekstraksi fitur dilakukan untuk mengubah data sensor yang bersifat deret waktu menjadi karakteristik numerik yang merepresentasikan pola perubahan suhu dan kelembapan pada setiap trial. Fitur yang digunakan adalah Temperature mean (°C), Temperature mean, Temperature slope, Humidity mean (%), Humidity min (%), Humidity slope. Dengan demikian, setiap trial direpresentasikan menggunakan fitur yang digunakan dan juga fitur slope digunakan untuk menggambarkan kecenderungan perubahan sensor selama proses pengeringan, sedangkan fitur statistik lainnya digunakan untuk menggambarkan karakteristik tingkat dan variasi suhu maupun kelembapan.

2.5. Clustering

Clustering dilakukan untuk mengidentifikasi kelompok pola proses pengeringan berdasarkan karakteristik suhu dan kelembapan. Algoritma K-Means digunakan karena mampu mengelompokkan data berdasarkan kedekatan karakteristik antar-trial dan relatif sederhana untuk diterapkan pada dataset berukuran kecil. Sebelum proses clustering, fitur dinormalisasi menggunakan StandardScaler agar setiap fitur memiliki skala yang sebanding. Jumlah cluster ditentukan dengan membandingkan nilai inertia menggunakan Elbow Method dan kualitas pemisahan kelompok menggunakan Silhouette Score. Pengujian dilakukan pada beberapa kandidat jumlah cluster, yaitu K=2 hingga K=6. Berdasarkan hasil evaluasi, K=2 menghasilkan Silhouette Score sebesar 0,2852, sehingga digunakan sebagai konfigurasi clustering dalam analisis selanjutnya. Hasil clustering kemudian digunakan untuk membentuk dua kelompok karakteristik proses pengeringan.

2.6. Analisis dan Interpretasi

Analisis dilakukan dengan membandingkan karakteristik fitur suhu dan kelembapan pada masing-masing cluster. Perbandingan dilakukan berdasarkan nilai rata-rata fitur untuk mengetahui karakteristik utama setiap kelompok. Selain itu, perbedaan antar-cluster dianalisis berdasarkan fitur yang memiliki selisih nilai paling besar.

HASIL DAN PEMBAHASAN

3.1. Karakteristik Dataset

Dataset yang digunakan merupakan data hasil pemantauan proses pengeringan berbasis sensor IoT. Data terdiri atas 30 trial (T01–T30) dengan jumlah observasi yang berbeda pada setiap trial, yaitu antara 31–36 observasi, sehingga total terdapat 1.004 observasi. Karakteristik data utama ditunjukkan pada Tabel 1.

Tabel 1. Karakteristik Dataset

Variable	Min	Max	Rata-Rata	Std
Temperature (°C)	28,21	39,00	35,027	2,196
Humidity (%RH)	40,39	79,11	58,763	7,608
Initial Weight (g)	453,87	622,70	538,246	51,058
Current Weight (g)	384,04	622,74	502,250	53,997
Weight Loss (%)	-0,19	17,90	6,725	3,937

3.2. Pola Perubahan Suhu dan Kelembapan

Hasil visualisasi menunjukkan adanya pola perubahan suhu dan kelembapan selama proses pengeringan. Suhu secara umum mengalami peningkatan dari sekitar 30,1 °C pada awal proses menjadi sekitar 37 °C pada akhir proses. Sebaliknya, kelembapan menunjukkan kecenderungan menurun dari sekitar 71% menjadi sekitar 48–51%. Pola tersebut menunjukkan hubungan yang berlawanan antara suhu dan kelembapan selama proses pengeringan. Ketika suhu meningkat, kelembapan cenderung mengalami penurunan. Kondisi ini juga terlihat pada hasil korelasi, di mana suhu dan kelembapan memiliki korelasi sebesar -0,692. Temuan tersebut memperlihatkan bahwa kedua parameter sensor memiliki informasi yang saling melengkapi dalam menggambarkan kondisi proses. Suhu merepresentasikan kecenderungan peningkatan kondisi termal, sedangkan kelembapan menunjukkan perubahan kondisi udara selama proses pengeringan.

3.3. Ekstraksi Karakteristik Setiap Trial

Untuk melakukan clustering, data time-series pada setiap trial diubah menjadi karakteristik statistik. Fitur yang digunakan meliputi: (1) temperature_mean, (2) temperature_std, (3) temperature_slope, (4) humidity_mean, (5) humidity_min, (6)

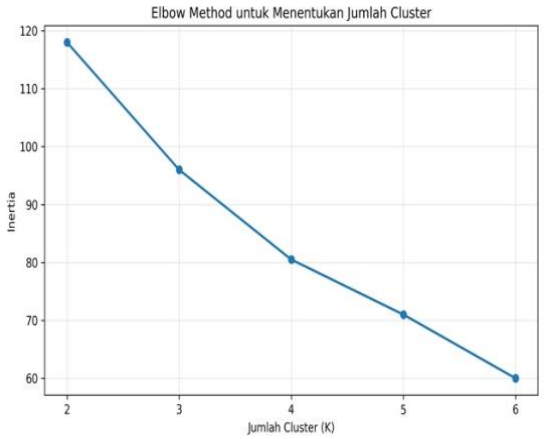
humidity_slope. Pemilihan fitur tersebut bertujuan mempertahankan informasi utama dari data sensor tanpa menggunakan seluruh titik pengamatan. Nilai mean menggambarkan kondisi rata-rata selama proses, minimum menggambarkan kondisi terendah yang terjadi, standard deviation menunjukkan tingkat variasi, sedangkan slope menggambarkan kecenderungan perubahan parameter terhadap waktu. Pendekatan ini menghasilkan 30 profil trial, yang kemudian menjadi unit analisis dalam proses clustering.

3.4. Hubungan Antarfitur

Analisis korelasi menunjukkan adanya hubungan yang cukup kuat antarfitur suhu maupun antarfitur kelembapan. temperature_mean memiliki korelasi sangat tinggi dengan temperature_min (r = 0,98) dan temperature_max (r = 0,99). Pada kelompok kelembapan, humidity_mean memiliki korelasi sebesar 0,97 dengan humidity_max, sedangkan humidity_std dan humidity_slope memiliki korelasi negatif yang sangat kuat (r = -0,97). Temuan tersebut menunjukkan bahwa karakteristik suhu dan kelembapan memiliki hubungan internal yang kuat. Oleh karena itu, penggunaan beberapa fitur yang mewakili aspek berbeda tingkat, variasi, dan kecenderungan perubahan lebih relevan dibandingkan hanya menggunakan satu nilai pengukuran. Korelasi antara parameter sensor dengan weight loss juga menunjukkan hubungan yang cukup kuat, yaitu 0,756 antara suhu dan weight loss, serta -0,846 antara kelembapan dan weight loss. Namun, variabel weight loss tidak digunakan sebagai fitur clustering, karena penelitian difokuskan pada pembentukan kelompok berdasarkan karakteristik suhu dan kelembapan.

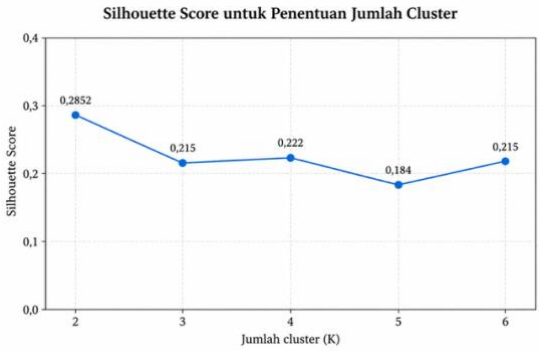
3.5. Penentuan Jumlah Cluster

Penentuan jumlah cluster dilakukan menggunakan Elbow Method dan Silhouette Score dengan kandidat jumlah cluster (K=2) hingga (K=6). Hasil Elbow Method menunjukkan nilai inertia pada Gambar 2.



Gambar 2. Elbow Method

Penurunan inertia semakin kecil ketika jumlah cluster bertambah. Hal tersebut menunjukkan bahwa penambahan cluster setelah beberapa titik tidak lagi memberikan penurunan inertia sebesar pada tahap awal. Sementara itu, hasil Silhouette Score ditunjukan pada Gambar 3.



Gambar 3. Silhouette Score

Nilai tertinggi diperoleh pada K=2 dengan Silhouette Score 0,2852. Berdasarkan kombinasi evaluasi tersebut, penelitian menetapkan dua cluster sebagai konfigurasi utama. Nilai Silhouette 0,2852 menunjukkan bahwa struktur cluster yang terbentuk belum memiliki pemisahan yang sangat kuat, tetapi terdapat struktur kelompok yang dapat digunakan untuk eksplorasi karakteristik trial. Hal ini penting untuk disampaikan secara transparan karena tujuan penelitian bukan membangun model prediksi, melainkan menemukan pola kelompok pada data sensor.

3.6. Profil Cluster

Hasil K-Means menghasilkan dua kelompok trial dengan karakteristik suhu dan kelembapan yang berbeda. Detail hasil ditunjukan pada Tabel 2.

Tabel 2. Profil Cluster

Fitur	Cluster 1	Cluster 2
Temperature mean (°C)	34,673	35,731
Temperature mean	1,904	1,954
Temperature slope	0,174	0,193
Humidity mean (%)	58,101	60,163
Humidity min (%)	45,484	43,459
Humidity slope	-0,668	-0,872

Cluster 2 memiliki rata-rata suhu yang lebih tinggi dibandingkan Cluster 1, yaitu 35,731 °C dibandingkan 34,673 °C. Cluster 2 juga memiliki temperature slope yang lebih tinggi, menunjukkan kecenderungan peningkatan suhu yang sedikit lebih besar selama proses. Pada parameter kelembapan, Cluster 2 memiliki rata-rata kelembapan sebesar 60,163%, lebih tinggi dibandingkan Cluster 1 sebesar 58,101%. Namun, nilai minimum kelembapan Cluster 2 lebih rendah, yaitu 43,459%, dibandingkan 45,484% pada Cluster 1. Perbedaan paling menarik terlihat pada humidity slope: Cluster 1 sebesar -0,668, sedangkan Cluster 2 sebesar -0,872. Artinya, Cluster 2 menunjukkan kecenderungan penurunan kelembapan yang lebih cepat selama proses pengeringan. Dengan demikian, kedua cluster dapat diinterpretasikan sebagai dua profil dinamika lingkungan pengeringan. Cluster 1 cenderung memiliki peningkatan suhu dan penurunan kelembapan yang relatif lebih moderat, sedangkan Cluster 2 menunjukkan suhu rata-rata yang lebih tinggi dan perubahan kelembapan yang lebih cepat.

3.7. Analisis

Hasil clustering menunjukkan bahwa data sensor IoT mampu membentuk kelompok berdasarkan karakteristik dinamika suhu dan kelembapan antartrial. Temuan ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa data sensor suhu dan kelembapan dapat digunakan sebagai dasar analisis berbasis machine learning, termasuk clustering untuk menemukan pola kondisi lingkungan. Pada konteks pengeringan, Pinheiro et al. (2023) juga menunjukkan bahwa metode unsupervised machine learning,

termasuk K-Means, dapat digunakan untuk mengelompokkan karakteristik proses pengeringan. Yang menjadi poin penting dari penelitian ini adalah unit analisis berupa trial proses pengeringan, bukan sekadar titik pembacaan sensor. Dengan melakukan ekstraksi fitur statistik dan temporal, 1.004 observasi sensor diringkas menjadi 30 profil trial sehingga pola antarpercobaan dapat dianalisis secara lebih sederhana. Hasil dua cluster memperlihatkan bahwa perbedaan kelompok terutama berkaitan dengan tingkat suhu rata-rata dan laju perubahan kelembapan. Nilai Silhouette sebesar 0,2852 juga memberikan informasi penting mengenai kualitas struktur kelompok. Cluster yang terbentuk dapat digunakan sebagai profil eksploratif, tetapi belum tepat diklaim sebagai kelompok yang memiliki batas pemisahan sangat kuat. Dengan demikian, hasil penelitian lebih tepat diposisikan sebagai identifikasi pola karakteristik proses pengeringan berbasis data sensor IoT daripada sebagai sistem pengambilan keputusan otomatis.

SIMPULAN

Berdasarkan hasil penelitian, data proses pengeringan dapat dikelompokkan berdasarkan karakteristik suhu dan kelembapan menggunakan metode K-Means Clustering. Hasil evaluasi jumlah cluster menggunakan Elbow Method dan Silhouette Score menunjukkan bahwa K=2 merupakan jumlah cluster yang paling sesuai, dengan Silhouette Score tertinggi sebesar 0,2852. Hal ini menunjukkan adanya dua pola karakteristik kondisi lingkungan selama proses pengeringan. Hasil analisis profil cluster menunjukkan bahwa kelembapan merupakan karakteristik yang paling membedakan kedua cluster, terutama pada nilai humidity mean dan humidity minimum. Cluster 1 memiliki rata-rata kelembapan sebesar 58,101%, sedangkan Cluster 2 sebesar 60,163%. Sementara itu, karakteristik suhu menunjukkan perbedaan yang lebih kecil, dengan rata-rata suhu masing-masing 34,673°C dan 35,731°C. Dengan demikian, clustering berhasil mengidentifikasi dua pola kondisi lingkungan pada proses pengeringan, dengan kelembapan sebagai karakteristik pembeda utama dan suhu sebagai karakteristik pendukung. Hasil ini dapat menjadi dasar untuk analisis lebih lanjut mengenai pola kondisi lingkungan pada proses pengeringan tanpa memerlukan tahap klasifikasi.

DAFTAR PUSTAKA

Dooyum Uyeh, D., Akinsoji, A., Asem-Hiablie, S., Itoro Bassey, B., Osinuga, A., Mallipeddi, R., Amaizu, M., Ha, Y., & Park, T. (2022). An online machine

learning-based sensors clustering system for efficient and cost-effective environmental monitoring in controlled environment agriculture. Computers and Electronics in Agriculture, 199, 107139. https://doi.org/10.1016/j.compag.2022.107139

Galih Pradana, M., Palilingan, K., Vanli Akay, Y., Puspasari Wijaya, D., & Hari Saputro, P. (2023). Comparison of Multi Layer Perceptron, Random Forest & Logistic Regression on Students Performance Test. 462–466. https://doi.org/10.1109/icimcis56303.2022.10017501

Gudivaka, R. L., Alabdeli, H., Sunil Kumar, V., Sushama, C., & Muthu, B. (2024). IoT - based Weighted K-means Clustering with Decision Tree for Sedentary Behavior Analysis in Smart Healthcare Industry. 2024 Second International Conference on Data Science and Information System (ICDSIS), 1–5. https://doi.org/10.1109/ICDSIS61070.2024.10594075

Jasmine, A., & Triawati, E. (2022). IoT-based system for monitoring the drying time of date seeds in the manufacturing of date coffee. Applied Research and Smart Technology (ARSTech), 3(1), 1–8. https://doi.org/10.23917/arstech.v3i1.480

Kumar, G. K., Bangare, M. L., Bangare, P. M., Kumar, C. R., Raj, R., Arias-Gonzáles, J. L., Omarov, B., & Mia, Md. S. (2024). Internet of things sensors and support vector machine integrated intelligent irrigation system for agriculture industry. Discover Sustainability, 5(1), 6. https://doi.org/10.1007/s43621-024-00179-5

Liu, H., & Qiu, X. (2026). Machine learning for IoT sensor data: enhancing environmental sensing with AI algorithms. Discover Artificial Intelligence, 6(1), 734. https://doi.org/10.1007/s44163-025-00476-6

Ngo, Q.-D., & Phan, L.-H. (2025). Analysis of Multivariate Indoor Building Data: A Comparative Study of Time-Series Clustering Methods. IEEE Access, 13, 139018–139032. https://doi.org/10.1109/ACCESS.2025.3596297

Novando, K., & P, N. (2025). Implementation of IoT-Based Automatic Irrigation System Using Decision Tree Algorithm on Hydroponic Garden at Institut Shanti Bhuana Bengkayang. Journal of Computing Innovations and Emerging Technologies, 1(2), 26–32. https://doi.org/10.64472/jciet.v1i2.6

Pal, A., & Gope, A. (2024). Texture identification in liquid crystal-protein droplets using evaporative drying, generalized additive modeling, and K-means Clustering. The European Physical Journal E, 47(5), 35. https://doi.org/10.1140/epje/s10189-024-00429-4

Pinheiro, R. de M., Gadotti, G. I., Bernardy, R., Tim, R. R., Pinto, K. V. A., & Buck, G. (2023). Computer vision by unsupervised machine learning in seed drying process. Ciência e Agrotecnologia, 47. https://doi.org/10.1590/1413-7054202347018922

Rami, R., & Malek, Dr. Z. (2024). Hybrid Approach For Anomaly Detection Using Clustering Mechanism. Educational Administration Theory and Practices. https://doi.org/10.53555/kuey.v30i5.5095

Utami, A., & Hadi, N. T. (2025). Anomaly Detection of Road Ranking Shifts Due to Traffic Accidents Using Deep Learning on Time Series Data. Journal of

Computing Innovations and Emerging Technologies, 1(1), 21–25. https://doi.org/10.64472/jciet.v1i1.5

W. Hamaali, G., A. Abduljabbar, K., & R. Sulaiman, Diary. (2023). K-means Clustering and PSO Algorithm for Wireless Sensor Networks Optimization. University of Thi-Qar Journal for Engineering Sciences, 13(1), 40–50. https://doi.org/10.31663/tqujes13.1.457(2023)

Waworundeng, J. (2024). IoT-based Environmental Monitoring with Data Analysis of Temperature, Humidity, and Air Quality. CogITo Smart Journal, 10(1), 271–284. https://doi.org/10.31154/cogito.v10i1.708.692-705

Xing, S., Lin, Z., Gao, X., Wang, D., Liu, G., Cao, Y., & Liu, Y. (2024). Research on Outgoing Moisture Content Prediction Models of Corn Drying Process Based on Sensitive Variables. Applied Sciences, 14(13), 5680. https://doi.org/10.3390/app14135680