

Pemetaan Kondisi Lingkungan Tanam Menggunakan K-Means Clustering

Diterima:
18 Februari 2023
Revisi:
10 Maret 2023
Terbit:
15 Maret 2023

^{1*}Ni'ma Kholila, ²M. Mujiono, ³Dona Wahyudi
¹⁻³Akademi Komunitas Negeri Putra Sang Fajar Blitar

Abstrak—Melalui pendekatan clustering, data rekam yang terdiri dari waktu tanam, pH, suhu air, suhu udara, dan nilai TDS dapat dimanfaatkan guna memetakan kondisi lingkungan tanam. Melalui pemetaan kondisi lingkungan tanam, petani mendapatkan informasi tambahan terkait kondisi lingkungan tanam milik petani. Melalui informasi tersebut, diharapkan petani dapat memberikan tindakan pertanian yang sesuai dengan kondisi lingkungan tanam, lebih efektif dan efisien. Pemetaan kondisi lingkungan tanam menggunakan k-means clustering algorithm dalam 3 (tiga) cluster. Masing-masing merupakan cluster kondisi lingkungan tanam yang kurang nutrisi dan kurang air, cukup nutrisi tetapi kurang air, serta cukup nutrisi dan cukup air. Cluster testing menggunakan elbow method menunjukkan bahwa jumlah cluster optimal dalam pengelompokan K-Means adalah 3 cluster dengan nilai inersia 199.065.

Kata Kunci— Pemetaan; Kondisi Lingkungan Tanam K-Means

Abstract—Through the clustering approach, recorded data consisting of planting time, pH, water temperature, air temperature, and TDS can be used to map agricultural environmental conditions. Through Mapping the Conditions of the Planting Environment, farmers get additional information regarding the conditions of the farmers' planting environment. Through this information, it is hoped that farmers can provide agricultural actions that are by the conditions of the planting environment, more effectively and efficiently. Mapping the conditions of the planting environment using the k-means clustering algorithm in 3 (three) clusters. Each of them is a cluster of growing environmental conditions that lack nutrition and lack of water, sufficient nutrition but lack of water, as well as sufficient nutrition and sufficient water. Cluster testing using the elbow method shows that the optimal number of clusters in the K-Means clustering is 3 clusters with an inertia value of 199,065.

Keywords— *Mapping, Agricultural Environmental Conditions, K-Means*

This is an open access article under the CC BY-SA License.



Penulis Korespondensi:

Ni'ma Kholila,
Administrasi Server dan Jaringan Komputer,
Akademi Komunitas Negeri Putra Sang Fajar Blitar,
Email: lila@akb.ac.id

I. PENDAHULUAN

Data BPS menunjukkan persentase angkatan kerja pertanian tahun 2021 mengalami peningkatan sebesar 10.74% dari tahun 2020. Peningkatan tersebut tentu tidak lebih besar dari peningkatan persentase konsumsi produk pertanian pada tahun 2021 sebesar 12.89% dari tahun 2020. Hal ini diperparah dengan persentase peningkatan luasan lahan pertanian pada tahun 2021 sebesar 1.87% dibandingkan luasan lahan pertanian pada tahun 2020[1]. Penggunaan *Internet of Things* (IoT) di bidang pertanian merupakan salah satu bentuk modernisasi di bidang pertanian. Kemampuan penggunaan internet sebagai penghubung berbagai perangkat, berbagai jenis sensor, dan objek cerdas lain, ke perangkat seluler, dapat membantu kegiatan pertanian lebih efektif dan efisien. Proses pemantauan perkembangan tanaman, kelembaban tanah, dan hama tanaman dari jarak jauh menjadi sangat mungkin menggunakan *Internet of Things* (IoT) [2].

Data rekam yang diperoleh sebagai akibat instalasi *Internet of Things* (IoT) sangat mungkin dikembangkan untuk mendapatkan manfaat yang lebih luas. Data rekam pertanian yang terkumpul dalam jumlah yang besar, dapat dimanfaatkan dengan pendekatan data mining untuk memperoleh informasi yang peneliti butuhkan, tentu untuk mendukung kegiatan-kegiatan pertanian lain. Pada penelitian terdahulu, produktivitas pertanian dipetakan menggunakan teknik pengelompokan data mining sehingga diketahui kelompok-kelompok lahan pertanian dengan produktivitas melebihi target, sesuai target, dan kurang dari target. Pengelompokan dilakukan berdasarkan data persentase peningkatan produksi pertanian [3]. Penelitian lain mengelompokkan lahan pertanian dalam 2 *cluster* berdasarkan tingkat kesuburan tanah yang dilakukan berdasarkan asesmen terhadap kandungan nutrisi yang terkandung dalam tanah [4]. *K-Means clustering* pada penelitian lain digunakan dalam melakukan clustering daerah tanam sehingga terbagi dalam 3 *cluster*, yaitu dataran tinggi, dataran rendah, dan pesisir pantai untuk selanjutnya mempengaruhi rekomendasi penanaman jenis sayur. [5]. *K-Means* sendiri merupakan pengelompokan klasik yang umum digunakan [6]. *K-Means algorithm* rentan terhadap *outlier*. Sehingga *outlier* harus dihilangkan sehingga dapat meningkatkan efisiensi *K-means algorithm* secara signifikan [7].

Berdasarkan uraian tersebut, peneliti memanfaatkan data rekam yang tersedia sehingga diperoleh informasi berupa Pemetaan Kondisi Lingkungan Tanam menggunakan *K-Means Clustering* dengan memanfaatkan data rekam pertanian yang diperoleh dari instalasi *Internet of Things* (IoT) di CV. Agro Utama Mandiri Lestari.

II. METODE

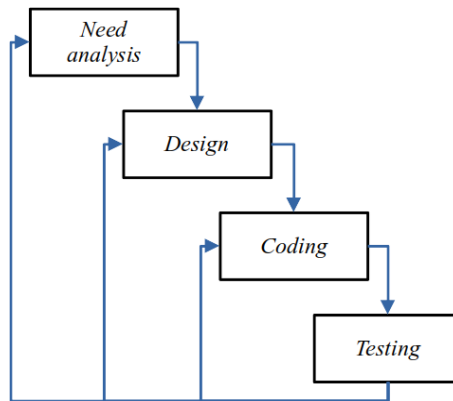
Pada tahap analisis kebutuhan, peneliti melakukan studi literatur terkait pengetahuan dan informasi yang berkaitan dengan kondisi lingkungan tanam. Berdasarkan penelitian terdahulu yang berhasil penulis himpun, terdapat beberapa penelitian yang memanfaatkan kondisi lingkungan tanam untuk berbagai kebutuhan. Tabel 1 menunjukkan pengelolaan data kondisi lingkungan tanam untuk membantu petani.

Tabel 1. Penelitian Terdahulu

Peneliti	Judul Penelitian	Hasil Penelitian
W. W. Widiyanto, F. Nugroho, Kusri (2019)	Implementation of the K-Means Cluster Algorithm in Rice Production Mapping and as a Decision Support for Agricultural Function Transition	Clustering menghasilkan kelompok produktivitas lahan mulai dari tinggi, sedang, dan tidak produktif sebagai acuan rekomendasi produksi pertanian hingga pengalihan fungsi lahan.
H. I. Hayatu, A. Mohammed, A. B. Isma'eel, Y. S. Ali (2020)	K-Means Clustering Algorithm based Classification of Soil Fertility in North West Nigeria	Penelitian ini mengelompokkan lahan pertanian dalam 2 cluster berdasarkan tingkat kesuburan tanah. Rekomendasi produksi pertanian dilakukan berdasarkan asesmen terhadap kandungan nutrisi yang terkandung dalam tanah.
S. Vaishnavi, M. Shobana, R. Sabitha, S. Karthik (2021)	Agricultural Crop Recommendations based on Productivity and Season	Hasil penelitian berupa rekomendasi jenis tanaman berdasarkan produktivitas lahan dan musim.
N. B. Nugraha, E. Alimudin, B. V. Indriyono (2022)	Implementasi K-Means Clustering Pada Sistem Pakar Penentuan Jenis Sayuran	Penelitian menghasilkan 3 cluster daerah yakni dataran tinggi, dataran rendah, dan pesisir pantai untuk selanjutnya mempengaruhi rekomendasi penanaman jenis sayur.

Berdasarkan penelitian terdahulu, terdapat beberapa kondisi lingkungan tanam yang membutuhkan pengamatan khusus dan secara signifikan mempengaruhi kondisi tanaman. Beberapa nilai kondisi lingkungan tanam yang perlu pengamatan khusus yaitu, waktu tanam, ph air, suhu air, suhu udara, dan nilai TDS merupakan komponen lingkungan tanam yang mampu menggambarkan kondisi lingkungan tanam sebagaimana hasil monitoring yang dilakukan oleh IoT-Agri [8].

Pada penelitian ini peneliti menggunakan model waterfall sebagai metodologi penelitian sebagaimana ditunjukkan pada Gambar 1 berikut:



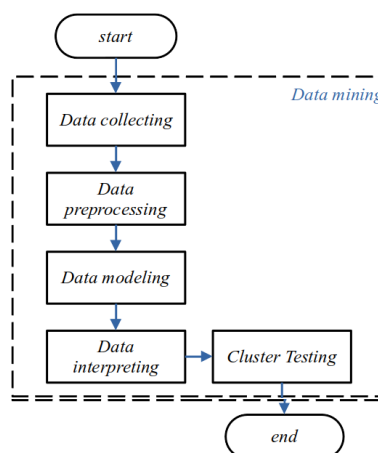
Gambar 1. *Waterfall*

Model *waterfall* merupakan model pengembangan klasik yang memiliki alur atau tahapan yang saling berurutan atau sekuensial. Alur atau tahapan yang cenderung sederhana menjadikan model *waterfall* banyak digunakan dalam kegiatan pengembangan [9]–[12].

A. Analisis Kebutuhan

Instalasi alat *Internet of Things* (IoT) di CV. Agro Utama Mandiri Lestari menghasilkan data rekam pertanian dalam jumlah besar. Data rekam belum dimanfaatkan secara maksimal untuk menunjang optimalisasi kegiatan pertanian. Data rekam meliputi data *timestamp*, ph, suhu air, suhu udara, dan nilai TDS. Nilai ph, suhu air, suhu udara, dan TDS cukup menggambarkan dan merepresentasikan kondisi lingkungan tanam [13]. Data rekam bersifat abstrak, belum terdapat pola (*pattern*) yang jelas. Data clustering dibutuhkan untuk memetakan kondisi lahan pertanian. Kondisi lahan pertanian yang telah terpetakan akan digunakan untuk menyusun rekomendasi tindakan pertanian. Sehingga selaras dengan tujuan instalasi IoT, tindakan pertanian lebih terukur dan kegiatan pertanian menjadi lebih optimal.

B. Desain



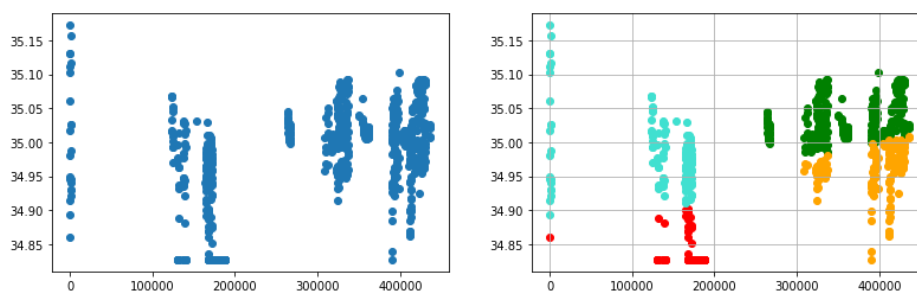
Gambar 2. Diagram Proses Kerja

Gambar 2 menunjukkan diagram proses kerja yang dimulai dengan tahap *data collecting*, pengumpulan seluruh data rekam dari CV. Agro Utama Mandiri Lestari. Tahap selanjutnya adalah

data preprocessing, meliputi *data cleansing* dan *data scaling*. *Data cleansing* bertujuan untuk menghilangkan data *outlier* serta data yang memiliki *missing value* [14]. *Data scaling* dilakukan untuk mentransformasikan dataset pada skala tertentu. Setelah melalui tahap data preprocessing, maka data siap digunakan dalam tahap berikutnya. Tahap berikutnya adalah *data modeling* menggunakan *K-Means algorithm*. *Data modeling* ditujukan untuk memetakan data rekam dalam *cluster* kondisi lingkungan tanam. Melalui *data modeling*, data dengan similaritas paling tinggi akan menempati *cluster* yang sama, demikian sebaliknya. Data dengan similaritas rendah akan ditempatkan pada *cluster* yang lain. Hasil pemetaan (*clustering*) akan diinterpretasikan pada tahap *data interpreting*. Masing-masing *cluster* akan dikenali, diidentifikasi, dan diinterpretasikan sebagai *cluster A*, *cluster B*, dan seterusnya. Tahap terakhir adalah *cluster testing*. Sebagai tahap penutup dalam pendekatan *data mining*, *clustering testing* digunakan untuk memastikan seberapa optimal *K-Means algorithm* mengelompokkan data dalam jumlah cluster tertentu. Melalui *cluster testing*, dapat diketahui jumlah *cluster* optimal yang sesuai dengan dataset yang peneliti miliki [15].

C. Data Mining

Sebagai bentuk teknik ekstraksi data, data mining memiliki kemampuan untuk menemukan pola (*pattern*) yang tersembunyi guna menghasilkan pengetahuan baru dalam suatu kumpulan data yang bahkan abstrak sebelumnya [16], [17]. Secara khusus pendekatan data mining memiliki teknik-teknik tertentu berdasarkan tujuan pemanfaatan ekstraksi data, baik untuk kebutuhan estimasi, prediksi, klasifikasi, pengelompokan, maupun asosiasi [18]. Teknik data mining tidak hanya disesuaikan dengan tujuan pemanfaatan ekstraksi data, tetapi juga disesuaikan dengan kondisi data yang hendak diekstrak. *Clustering* merupakan salah satu bentuk teknik data mining yang dapat dimanfaatkan untuk pengelompokan atau pemetaan. *Clustering* melakukan pengelompokan data berdasarkan kemiripan data sehingga terdapat kelompok dengan kumpulan data yang mirip dengan data lain dalam satu kelompok, tetapi tidak mirip dengan data di kelompok lain [3]. Gambar 3 menunjukkan ilustrasi data sebelum dan sesudah menggunakan teknik *clustering* dengan *K-Means algorithm*.



Gambar 3. Ilustrasi Data Sebelum dan Sesudah Clustering

Langkah pertama yang dilakukan adalah menentukan berapa banyak *cluster* yang akan dibuat, selanjutnya disebut sebagai nilai k . Tahap selanjutnya adalah menentukan nilai *centroid* (pusat *cluster* awal) yang diperoleh dari data set secara acak sejumlah nilai k . Selanjutnya menghitung jarak setiap dataset dengan masing-masing *centroid*. Perhitungan menggunakan *euclidean distance*, proses perhitungan jarak dilakukan untuk menentukan jarak terdekat dari seluruh data dengan masing-masing *centroid* [3]. Adapun persamaan *euclidean distance* ditunjukkan oleh persamaan (1):

$$d(x_i, \mu_i) = \sqrt{(x_i - \mu_i)^2} \quad (1)$$

x_i : nilai dataset

μ_i : nilai *centroid*

Selanjutnya merupakan proses pengelompokan data berdasarkan nilai *euclidean distance* terkecil terhadap suatu *centroid* dalam suatu kelompok. Nilai *centroid* baru dihitung ulang berdasarkan nilai rata-rata kelompok [3], sebagaimana ditunjukkan oleh persamaan (2) sebagai berikut:

$$C_k = \frac{1}{n_k} \sum d \quad (2)$$

C_k : nilai *centroid* baru pada *cluster* k

n_k : jumlah data pada *cluster* k

d_i : dataset

Nilai *euclidean distance* hingga nilai *centroid* akan terus dihitung sampai nilai *centroid* tidak lagi berubah. Sehingga keanggotaan data pada suatu *cluster* juga tidak mengalami perubahan [3].

D. Cluster Testing

Pengujian *cluster* dilakukan untuk memastikan optimalnya hasil pemetaan (*clustering*) menggunakan *elbow method*. Ide dasar *elbow method* adalah menentukan jumlah *cluster* kemudian menambahkan *cluster*, menghitung *Sum Squared Error (SSE) per cluster* hingga jumlah maksimum *cluster* yang telah ditentukan, membandingkan perbedaan SSE setiap *cluster*, dan perbedaan paling ekstrim yang membentuk sudut siku menunjukkan jumlah *cluster* terbaik [19]. Adapun persamaan *euclidean distance* ditunjukkan oleh persamaan (3)[20]:

$$SSE = \sum_{K=1}^K \sum_{x_i} |x_i - c_k|^2 \quad (3)$$

K : *cluster* ke- c

x_i : jarak data obyek ke- i

c_k : pusat *cluster* ke- i

III. HASIL DAN PEMBAHASAN

Data rekam terdiri dari timestamp, ph, suhu air, suhu udara, dan nilai TDS. Kebun yang dimaksud berada dalam naungan CV. Agro Utama Mandiri Lestari, Jl. Raya Ngadiluwih, RT.03/RW.02, Munengan, Mangunrejo, Kec. Ngadiluwih, Kabupaten Kediri, Jawa Timur dengan rincian sebagaimana ditunjukkan pada Gambar 4.

```
<bound method DataFrame.count of          timestamp      ph  suhuair  suhuudara  tds
0      1666646665  34.983482 -127.000  28.139496  320
1      1666646670  34.988834 -127.000  28.085327   91
2      1666646676  35.136864 -127.000  28.032494  327
3      1666646682  34.911415 -127.000  27.982140  283
4      1666646688  35.044170 -127.000  27.948570  103
...
10409  1667077224  34.988003 -127.000  25.713730  236
10410  1669894208  34.919182 -127.000         NaN   95
10411  1667077224  34.988003 -127.000  25.713730  236
10412           0  17.413317 -127.000 -50.000000   0
10413  1833296485  35.194489   37.125  99.999817  521

[10414 rows x 5 columns]>
```

Gambar 4. Rincian Data Rekam

Data rekam selanjutnya melalui tahap *preprocessing* guna menghindari *outlier value* dan *missing value* menggunakan fungsi *isna()* dan *dropna()* pada *pandas python library*. Gambar 5 menunjukkan hasil *data preprocessing*:

```
<bound method DataFrame.count of          timestamp      ph  suhuair  suhuudara  tds
0      178446  34.827377 -127  26.525879   0
1      179545  34.827377 -127  26.525879   0
2      179657  34.827377 -127  26.525879   0
3      179769  34.827377 -127  26.526260   0
4      179117  34.827377 -127  26.526642   0
...
8858      344  35.137070 -127  27.317047  379
8859      206  35.074600 -127  27.467728  379
8860      350  35.130363 -127  27.313614  381
8861      357  35.129875 -127  27.315903  382
8862      225  35.038025 -127  27.448082  383

[8863 rows x 5 columns]>
```

Gambar 5. Hasil *Data Preprocessing*

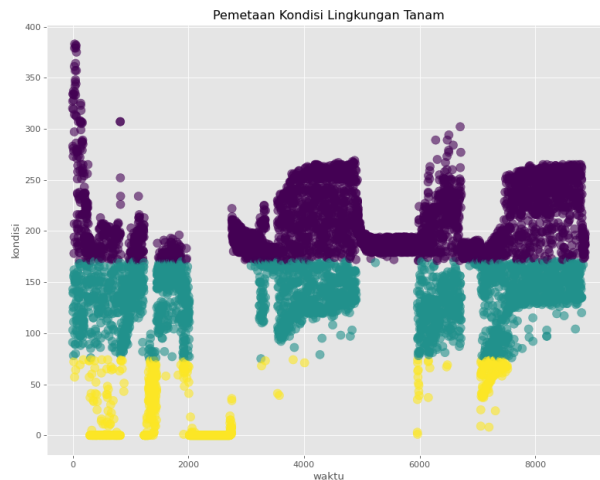
Hasil *preprocessing* siap digunakan dalam proses data modeling menggunakan fungsi *KMeans* pada *scikit-learn python library*. Pada penelitian ini, ditentukan 3 (tiga) nilai k dengan maksud membagi dataset dalam 3 (tiga) *cluster* kondisi lahan tanam. Gambar 6 menunjukkan hasil *data modeling*:

```
          timestamp      ph  suhuair  suhuudara  tds  cluster
0             1  34.983482 -127  28.139496  320      0
1             6  34.988834 -127  28.085327   91      1
2            12  35.136864 -127  28.032494  327      0
3            18  34.911415 -127  27.982140  283      0
4            24  35.044170 -127  27.948570  103      1
...
8858      436645  35.007305 -127  25.507736  187      0
8859      436651  35.021851 -127  25.510406  189      0
8860      436657  35.006390 -127  25.516129  179      0
8861      436663  35.005932 -127  25.511742  186      0
8862      436670  35.022755 -127  25.517273  197      0

[8863 rows x 6 columns]
```

Gambar 6. Hasil *Data Modeling*

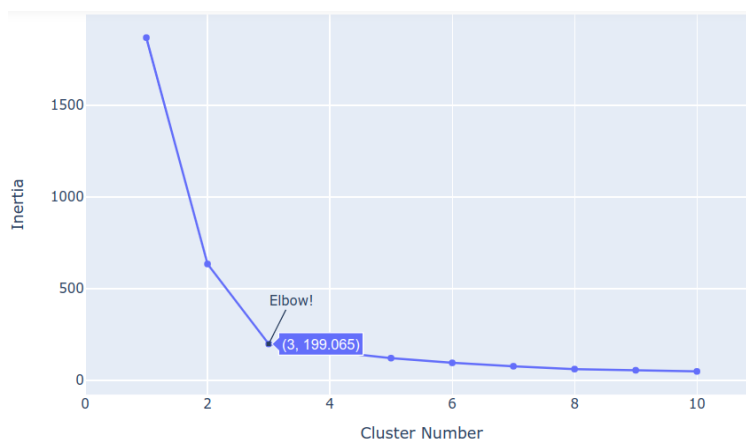
Gambar 7 merupakan visualisasi data berdasarkan kondisi lahan tanam yang melalui tahap *data modeling* menggunakan fungsi *pyplot* pada *matplotlib python library*.



Gambar 7. Pemetaan Kondisi Lingkungan Tanam

Berdasarkan visualisasi di atas, kondisi lingkungan tanam yang telah dipetakan dalam 3 (tiga) *cluster* dapat diinterpretasikan sebagai *cluster 1* (kuning) menunjukkan kondisi lingkungan tanam yang kekurangan nutrisi dan kekurangan air, *cluster 2* (biru) menunjukkan kondisi lingkungan tanam yang cukup nutrisi tetapi kekurangan air, dan *cluster 3* (biru) menunjukkan kondisi lingkungan tanam yang cukup nutrisi dan cukup air.

Cluster yang terbentuk pada tahap sebelumnya diuji menggunakan *elbow method* guna memastikan seberapa optimal *K-Means algorithm* melakukan klusterisasi dataset berdasarkan jumlah *cluster* yang ditentukan. Seberapa optimal *K-Means algorithm* ditunjukkan melalui nilai inersia yang diperoleh pada masing-masing percobaan jumlah *cluster*. Gambar 8 menunjukkan *cluster testing* menggunakan *elbow method* menggunakan fungsi *KMeans.inertia_* pada *scikit-learn python library*.



Gambar 8. *Elbow Method*

Gambar 8 menunjukkan pengujian dilakukan dalam 10 (sepuluh) skenario pengujian. Skenario pengujian pertama memetakan seluruh dataset dalam 1 (satu) cluster. Skenario pengujian kedua membagi seluruh dataset dalam 2 (dua) cluster. Skenario pengujian ketiga memetakan seluruh dataset dalam 3 (tiga) cluster, dan seterusnya. Pada masing-masing skenario pengujian dihitung nilai inersia. Jumlah cluster terbaik dapat dilihat dari nilai inersia tertinggi pada siku dataset diuji dalam 10 kondisi. Gambar 8 menampilkan nilai inersia pada masing-masing skenario pengujian (dalam *cluster*). Elbow point didapat dari poin yang terletak pada poin siku (*elbow*). Sehingga *elbow point* merujuk pada nilai inersia 199,065 pada percobaan pengujian dalam 3 (tiga) *cluster*. Melalui *cluster testing* menggunakan *elbow method*, maka dibenarkan bahwa nilai inersia terbaik terdapat pada pembagian 3 (*cluster*) sebagaimana peneliti lakukan pada tahap *data modeling*.

IV. KESIMPULAN

Peneliti menyimpulkan bahwa berdasarkan data rekam kondisi lahan tanam yang diperoleh melalui instalasi IoT-Agri di lahan tanam milik CV. Agro Utama Mandiri Lestari, dapat dipetakan kondisi lahan tanam dalam 3 (tiga) cluster. Cluster 1 menunjukkan kondisi lingkungan tanam yang kekurangan nutrisi dan kekurangan air. Cluster 2 menunjukkan kondisi lingkungan tanam yang cukup nutrisi tetapi kekurangan air. Cluster 3 menunjukkan kondisi lingkungan tanam yang cukup nutrisi dan cukup air. Penentuan jumlah cluster dikuatkan melalui cluster testing menggunakan elbow method dengan elbow point pada titik (3 , 199.065) yaitu nilai inersia 199,065 pada percobaan pengujian dalam 3 (tiga) cluster.

DAFTAR PUSTAKA

- [1] BPS, "Statistik Hortikultura 2021," vol. 21, no. 1, pp. 1–9, 2020.
- [2] A. Nalendra, M. N. Fuad, D. Wahyudi, N. Kholila, and M. Mujiono, "Effectiveness of the Use of the Internet of Things (IoT) in the Agricultural Sector," *International Journal of Science and Society*, vol. 4, no. 3, pp. 474–478, Sep. 2022, doi: 10.54783/ijssoc.v4i3.541.
- [3] W. W. Widiyanto, F. Nugroho, and K. Kusriani, "Implementation Of The K-Means Cluster Algorithm In Rice Production Mapping And As A Decision Support For Agricultural Function Transition," *Jurnal Informa : Jurnal Penelitian dan Pengabdian Masyarakat*, vol. 5, no. 4, pp. 46–52, 2020, doi: 10.46808/INFORMA.V5I4.155.
- [4] I. H. Hayatu, A. Mohammed, B. A. Ismael, and S. Y. Ali, "K-MEANS CLUSTERING ALGORITHM BASED CLASSIFICATION OF SOIL FERTILITY IN

- NORTH WEST NIGERIA,” FUDMA JOURNAL OF SCIENCES, vol. 4, no. 2, pp. 780–787, Nov. 2020, doi: 10.33003/fjs-2020-0402-363.
- [5] N. B. Nugraha, E. Alimudin, and B. V. Indriyono, “Implementasi K-Means Clustering Pada Sistem Pakar Penentuan Jenis Sayuran,” *Journal of Innovation Information Technology and Application (JINITA)*, vol. 4, no. 2, pp. 133–141, 2022, doi: 10.35970/JINITA.V4I2.1627.
- [6] B. Chong, “K-means clustering algorithm: a brief review,” *Academic Journal of Computing & Information Science*, vol. 4, no. 5, 2021, doi: 10.25236/ajcis.2021.040506.
- [7] K. P. Sinaga and M. S. Yang, “Unsupervised K-means clustering algorithm,” *IEEE Access*, vol. 8, pp. 80716–80727, 2020, doi: 10.1109/ACCESS.2020.2988796.
- [8] A. K. Nalendra, D. Wahvudi, M. Mujiono, M. N. Fuad, and N. Kholila, “IoT-Agri: IoT-based Environment Control and Monitoring System for Agriculture,” pp. 1–6, 2023, doi: 10.1109/ICIC56845.2022.10006964.
- [9] I. Yuniarfi and S. Saifulloh, “Penerapan Algoritma K-means Untuk Pengelompokan Usia Pasien Penerima Vaksin di KAB. Ngawi,” *Jurnal Informasi dan Komputer*, vol. 9, no. 2, pp. 51–62, 2021, doi: 10.35959/JIK.V9I2.211.
- [10] S. A. B. C. Busro et al., “Rancangan Pembuatan API Website Data Tanaman Obat Dan Langka Kabupaten Kediri,” *Bulletin of Information Technology (BIT)*, vol. 3, no. 4, pp. 255–260, Dec. 2022, doi: 10.47065/BIT.V3I4.373.
- [11] R. A. Purba and S. Sondang, “Design and Build Monitoring System for Pregnant Mothers and Newborns using the Waterfall Model,” *INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi*, vol. 6, no. 1, pp. 29–42, Feb. 2022, doi: 10.29407/INTENSIF.V6I1.16085.
- [12] I. Benawan, D. Mutiara, K. Nugraheni, B. Noranita, and G. Aryotejo, “Digital Education Game for TK-A Level Students Using Multimedia Development Life Cycle Method,” *INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi*, vol. 7, no. 1, pp. 68–83, Feb. 2023, doi: 10.29407/INTENSIF.V7I1.18671.
- [13] M. Iqbal, F. Barchia, and A. Romeida, “PERTUMBUHAN DAN HASIL TANAMAN MELON (*Cucumis melo* L.) PADA KOMPOSISI MEDIA TANAM DAN FREKUENSI PEMUPUKAN YANG BERBEDA,” *Jurnal Ilmu-Ilmu Pertanian Indonesia*, vol. 21, no. 2, pp. 108–114, 2019, doi: 10.31186/JIPI.21.2.108-114.
- [14] A. Susanto and ; Meiryani, “Functions, Processes, Stages And Application Of Data Mining,” *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, vol. 8, p. 7, 2019.

- [15] A. F. Febrianti, A. H. Cabral, and G. Anuraga, “K-Means Clustering Dengan Metode Elbow Untuk Pengelompokan Kabupaten Dan Kota Di Jawa Timur,” Seminar Nasional Hasil Riset dan Pengabdian -SNHRP, pp. 863–870, 2018.
- [16] S. Sunardi, A. Fadlil, and N. M. P. Kusuma, “Comparing Data Mining Classification for Online Fraud Victim Profile in Indonesia,” *INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi*, vol. 7, no. 1, pp. 1–17, Feb. 2023, doi: 10.29407/INTENSIF.V7I1.18283.
- [17] Sucipto, Kusrini, and E. L. Taufiq, “Classification method of multi-class on C4.5 algorithm for fish diseases,” in *Proceeding - 2016 2nd International Conference on Science in Information Technology, ICSITech 2016: Information Science for Green Society and Environment*, 2016, pp. 5–9. doi: 10.1109/ICSITech.2016.7852598.
- [18] M. A. Jassim and S. N. Abdulwahid, “Data Mining preparation: Process, Techniques and Major Issues in Data Analysis,” *IOP Conference Series: Materials Science and Engineering*, vol. 1090, no. 1, p. 12053, 2021, doi: 10.1088/1757-899X/1090/1/012053.
- [19] E. Umargono, J. E. Suseno, and V. G. S. K., “K-Means Clustering Optimization using the Elbow Method and Early Centroid Determination Based-on Mean and Median,” vol. 474, no. Isstec 2019, pp. 234–240, 2020, doi: 10.5220/0009908402340240.
- [20] D. Ayu, I. C. Dewi, and K. Pramita, “Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali,” *Matrix : Jurnal Manajemen Teknologi dan Informatika*, vol. 9, no. 3, pp. 102–109, Nov. 2019, doi: 10.31940/MATRIX.V9I3.1662.