

## CLUSTERING KABUPATEN BERDASARKAN LUAS HUTAN MENGGUNAKAN METODE K-MEANS DI PROVINSI JAWA TENGAH

### DISTRICT CLUSTERING BASED ON FOREST AREA USING K-MEANS METHOD IN CENTRAL JAVA PROVINCE

Yusri Eli Hotman Turnip<sup>1)</sup> dan Rina Fitriana<sup>1,2)\*</sup>

<sup>1)</sup>Magister Teknik Industri Universitas Trisakti, Jl. Kyai Tapa No.1, Grogol, Jakarta, Indonesia

<sup>2)</sup>Jurusan Teknik Industri Universitas Trisakti, Jl. Kyai Tapa No.1, Grogol, Jakarta, Indonesia

\*E-mail: [rinaf@trisakti.ac.id](mailto:rinaf@trisakti.ac.id)

Makalah: Diterima 16 Oktober 2022; Diperbaiki 2 Februari 2023; Disetujui 20 Februari 2023

#### ABSTRACT

Indonesia is one of the countries with the largest forest in the world. The tropical climate and high rainfall cause a lot of biodiversity in Indonesia's forests. The existence of these forests can be utilized by many parties, both the government and the community in accordance with their functions to improve welfare. The government through the Central Statistic Agency has provided data information related to the forest area in various regions, one of which is Central Java Province but still requires development to obtain important information in the data. This study aims to divide the district based on forest area including protected forest, protected area, area for production, and area for other users in Central Java province using the K-Means Data Mining method. The data is obtained from Central Statistic Agency for the Central Java area, where four types of forest are to be grouped. The results of this study indicate that the grouping of districts based on the area of forest owned is based on the smallest Davies Bouldin (DB), which is 0.436 in the grouping with 2 clusters. The two clusters are distinguished based on the value of the proximity of the forest type attribute with the centroid point in each cluster. The clustering process grouped 26 districts in the province of Central Java into cluster 1, while cluster 2 consisted of 3 districts in Central Java, namely Grobogan, Blora, and Brebes districts.

Keywords: clustering, forests, K-Means

#### ABSTRAK

Indonesia merupakan salah satu negara dengan hutan terluas di dunia. Iklim tropis dan curah hujan yang cukup tinggi, menyebabkan banyaknya keanekaragaman hayati di hutan-hutan Indonesia. Keberadaan hutan-hutan ini dapat dimanfaatkan banyak pihak baik pemerintah maupun masyarakat sesuai dengan fungsinya untuk meningkatkan kesejahteraan. Pemerintah melalui Badan Pusat Statistik telah menyediakan data informasi terkait luas hutan di berbagai daerah, salah satunya adalah provinsi Jawa Tengah, namun masih memerlukan pengembangan untuk mendapatkan informasi penting di dalam data tersebut. Penelitian ini bertujuan untuk membagi wilayah kabupaten berdasarkan luas area hutan antara lain hutan lindung, kawasan perlindungan, kawasan untuk produksi, dan kawasan untuk pengguna lain di provinsi Jawa Tengah menggunakan metode *Data Mining* yaitu *K-Means*. Data diperoleh dari Badan Pusat Statistik untuk area Jawa Tengah dimana terdapat empat jenis hutan yang akan dikelompokkan. Hasil dari penelitian ini menunjukkan pengelompokkan kabupaten berdasarkan luas hutan yang dimiliki berdasarkan indeks *Davies Bouldin* terkecil, yaitu sebesar 0,436 pada pengelompokan dengan 2 *cluster*. Kedua *cluster* dibedakan berdasarkan nilai jarak kedekatan *attribute* jenis hutan dengan titik *centroid* pada masing-masing *cluster*. Proses *clustering* yang dilakukan mengelompokkan 26 kabupaten di wilayah provinsi Jawa Tengah ke dalam *cluster* 1, sedangkan *cluster* 2 terdiri dari 3 kabupaten di wilayah Jawa Tengah yaitu kabupaten Grobogan, kabupaten Blora, dan kabupaten Brebes.

Kata kunci: clustering, hutan, K-Means

#### PENDAHULUAN

Ekosistem hutan menyediakan beragam barang dan fasilitas yang penting bagi kesejahteraan manusia. Namun masalah lingkungan global, peningkatan urbanisasi, tekanan industri, dan mekanisme pasar yang dominan menjadi ancaman ekosistem hutan. Pemanfaatan hutan umumnya berupa produk dengan bahan baku kayu, namun kurang memperhatikan aspek keberlangsungan

ekonomi untuk kepentingan masyarakat, seperti cadangan karbon, konservasi keanekaragaman hayati, atau nilai rekreasi (Maier *et al.*, 2021). Dalam pemanfaatan hutan tersebut, pengawasan terhadap perubahan area hutan dan pemicunya telah menjadi komponen penting dalam pengambilan keputusan dan pengembangan kebijakan. Hasil penelitian sebelumnya menunjukkan adanya penurunan keseluruhan dalam kategori luas hutan dan air dikawasan yang diteliti, sedangkan lahan pertanian,

padang rumput, dan lahan kosong meningkat (Rotich dan Ojwang, 2021). Pernyataan tersebut didukung oleh penelitian lain yang menyatakan terjadinya proses deforestasi di batas wilayah moratorium yang ditetapkan pemerintah. Hal ini diperkuat saat perusahaan yang memiliki hak konsesi dalam mengelola hutan melakukan kegiatan-kegiatan deforestasi yang mengakibatkan hilangnya hutan nasional (Leijten *et al.*, 2021). Saat ini sektor kehutanan memunculkan minat yang besar dalam bidang penelitian. Hal itu tidak lepas dari perhatian yang berkembang pada dimensi kelembagaan, kebijakan dan sosial khususnya dalam mendukung inovasi di sektor ini (Weiss *et al.*, 2021). Oleh karena itu, diperlukan penelitian-penelitian yang bisa memaksimalkan potensi hutan yang ada serta menjaganya tetap eksis demi keberlangsungan hidup manusia.

Hutan tropis memiliki suatu peranan penting dalam menjaga iklim di dunia dengan menjadi penyerap karbon dioksida terbesar (Alexander *et al.*, 2018). Dengan iklim tropis dan dilalui oleh garis khatulistiwa, Indonesia menjadi salah satu negara dengan hutan tropis yang sangat luas, selain didukung oleh faktor luasnya wilayah. Di dunia, luas hutan yang dimiliki di berbagai negara mencapai sekitar 20% atau sekitar 730 juta hektar dan dikelola oleh masyarakat setempat dengan berbagai skema seperti kepemilikan lahan, lisensi, dan manajemen (Rodd *et al.*, 2022). Peranan hutan dalam sistem produksi sendiri tidak hanya terbatas pada pengelolaan kayu atau lahan. Kecenderungan global saat ini terhadap bio-ekonomi telah mempengaruhi potensi peningkatan peran untuk produk hutan non-kayu (Muttalain dan Vilko, 2022).

Aktivitas *logging* di Indonesia diinisiasi oleh perusahaan besar sejak tahun 1970 yang diikuti oleh penerbitan izin untuk perusahaan kecil pada tahun 1998 namun disertai dengan kurangnya kontrol untuk manajemen praktis hutan. Sedangkan untuk proses pemulihan hutan setelah tebang pilih yang ditunjukkan dengan pemulihan stok karbon di permukaan tanah paling tidak 26 tahun dan berpotensi lebih lama tergantung pada intensitas penebangan (Rahayu *et al.*, 2022). Hal ini juga diperburuk dengan adanya kebakaran hutan dimana kebakaran hutan yang terjadi di banyak negara berkontribusi pada pemanasan global (Abdul Kadir *et al.*, 2022). Untuk itu perlu adanya regulasi akan legalitas pengelolaan hutan agar selain berguna bagi kesejahteraan masyarakat, kelestarian hutan di Indonesia dapat tetap terjaga dan dalam pengawasan pengambil kebijakan. Indonesia melalui kementerian lingkungan dan hutan menjadi regulator untuk sistem legalitas kayu (Maryudi *et al.*, 2021). Penelitian sebelumnya menyatakan bahwa pemberlakuan moratorium terbukti efektif untuk menghindari terjadinya proses deforestasi (Leijten *et al.*, 2021).

*Data mining* merupakan teknik dalam melakukan ekstraksi data yang berguna dari *big data*

untuk mendapatkan informasi atau pengetahuan dan termasuk dalam bagian metode *knowledge discovery in databases* (KDD) (Syadzali *et al.*, 2020) (Krishnamoorthy dan Karthikeyan, 2022). Penggunaan *data mining* dalam memperoleh suatu informasi dari suatu data telah banyak dilakukan pada penelitian terdahulu. Salah satunya yaitu penelitian untuk memperoleh informasi dari data pemasaran produk di pabrik roti menggunakan *data mining*. Penelitian ini mengelompokkan berbagai jenis produk yang dimiliki suatu pabrik roti berdasarkan jumlah produk yang tersisa dari beberapa cabang yang dimiliki (Fitriana *et al.*, 2018). Penelitian lain yang menerapkan penggunaan *data mining* yaitu pengelompokan jenis telur ayam *lehorn* dan *omega-3* berdasarkan data hasil pengolahan citra, karakteristik yang dihasilkan dari pengolahan citra tersebut selanjutnya dijadikan dasar pengelompokan telur kedalam kelompok *lehorn* atau *omega-3* (Nurhayati, 2020). *Data mining* menyediakan metodologi dan teknologi untuk mentransformasi data yang cukup besar menjadi informasi yang berguna untuk mengambil keputusan (Santos-Pereira *et al.*, 2021). Hal ini didukung oleh penelitian lainnya dimana *data mining* untuk mempermudah perolehan informasi yang mendalam (Fitriana *et al.*, 2017). *Data mining* memiliki hubungan dengan *artificial intelligence* (AI), *machine learning*, statistik, dan *database*. Beberapa metode yang cukup populer antara lain *neural network*, *classification*, *association rule*, *genetic algorithm*, dan *clustering* (Indraputra dan Fitriana, 2020).

*Clustering* merupakan salah satu algoritma yang ada pada *data mining*. Pendekatan yang dilakukan secara otomatis mengklasifikasi data kedalam beberapa kelompok obyek yang sama. *K-Means* adalah algoritma *clustering* iteratif yang menyempurnakan *cluster* yang ditugaskan melalui banyak iterasi. Tahap awal dilakukan dengan pemilihan jumlah *cluster* yang diinginkan atau dengan jumlah yang sama dengan *centroid*. Setiap titik data dialokasikan ke sebuah *cluster* berdasarkan jarak ke *centroid* dari semua *cluster* (Naghizadeh dan Metaxas, 2020). *Centroid* merupakan titik pusat yang didapatkan dari nilai rata-rata dari setiap *attribute* yang diteliti (Ida *et al.*, 2022). Penelitian lain mendefinisikan tujuan *K-Means* yaitu untuk mengelompokkan data berdasarkan input kriteria tertentu menjadi anggota kelompok (Rabiah *et al.*, 2015). *K-Means* merupakan salah satu algoritma *data mining* paling sering digunakan dan termasuk ke dalam jenis *unsupervised learning* (Ahmed *et al.*, 2020). *K-Means* menghasilkan partisi obyek yang jelas dengan menggunakan keanggotaan yang jelas pada domain Boolean 0 dan 1 (Ubukata *et al.*, 2021). Algoritma *K-Means* melakukan pengelompokan sebanyak 'k' kelompok dimana anggota kelompok dipilih berdasarkan jarak pada masing-masing data dan terhadap *centroid* (Rabiah *et al.*, 2015); (Pusadan *et al.*, 2019). Penerapan data *warehouse*

menggunakan model dimensional dengan skema bintang; pengelompokan wilayah produksi tanaman obat menggunakan algoritma K-Means (Hidayat dan Fitriana, 2022).

## METODE PENELITIAN

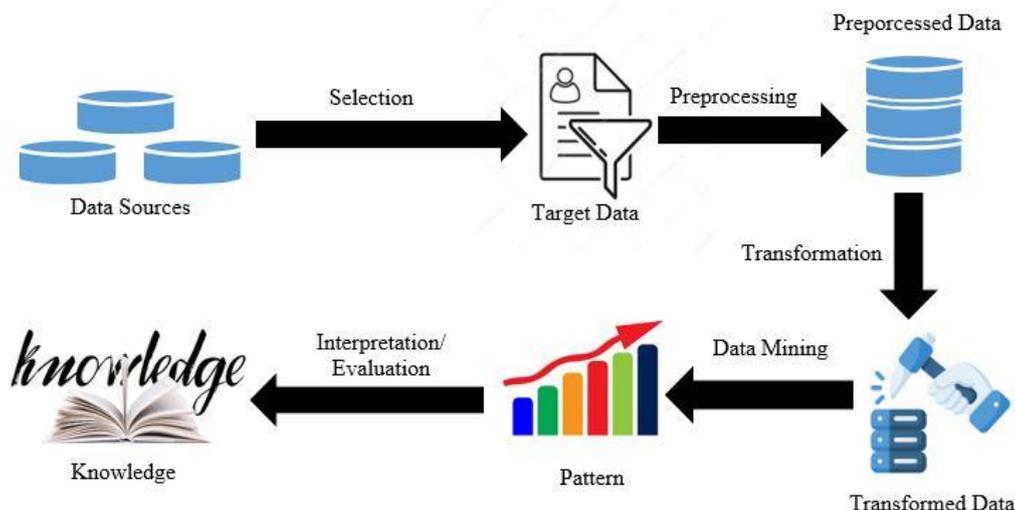
Pada penelitian ini, data yang digunakan merupakan data luas lahan hutan sesuai dengan jenis hutan di provinsi Jawa Tengah. Data yang didapatkan menginformasikan luas lahan untuk tahun 2019-2021 dimana data yang akan diambil hanya data untuk tahun 2021. Sumber data yang digunakan berasal dari Badan Pusat Statistik untuk wilayah provinsi Jawa Tengah pada website [www.jateng.bps.go.id](http://www.jateng.bps.go.id). Pengumpulan data dilakukan dengan mengambil data luas lahan berdasarkan jenis hutan melalui Badan Pusat Statistik untuk wilayah Jawa Tengah data yang diambil terdiri dari data luas lahan untuk 4 jenis hutan berdasarkan fungsinya, yaitu hutan lindung, kawasan perlindungan, kawasan untuk produksi, dan kawasan untuk pengguna lain. Pada informasi data yang didapatkan, luasan lahan dibagi untuk masing-masing kabupaten di provinsi Jawa Tengah. Masing-masing kabupaten ini selanjutnya akan dikelompokkan berdasarkan luas hutan yang dimiliki untuk masing-masing jenis hutan menggunakan metode *K-Means*.

Data yang digunakan merupakan data yang masuk ke dalam kategori pertanian dan pertambangan untuk bagian kehutanan dan telah diperbaharui pada tanggal 15 Maret 2022 (Badan Pusat Statistik, 2022). Kekurangan data yang dimunculkan yaitu data berupa tabel yang dapat diunduh dalam format *Ms. Excel* sehingga sulit untuk menarik suatu kesimpulan. Selain itu data yang disajikan tidak memiliki bentuk visual seperti grafik atau diagram yang mempermudah pembaca dalam menarik suatu kesimpulan. Untuk itu diperlukan visualisasi dari data-data yang dimiliki beserta pengolahan data untuk

melihat suatu pola yang muncul dari data-data tersebut, khususnya data *clustering* yang akan membagi kabupaten yang diteliti dengan tujuan pengawasan dan pengambilan keputusan oleh pihak terkait seperti pemerintah daerah maupun kementerian terkait. Dengan pengolahan data dan visualisasi yang dilakukan, keputusan yang diambil dapat tepat sasaran.

Gambar 1 memperlihatkan tahap pengolahan data informasi yang didapatkan menggunakan metode-metode *data mining* (Srivastava *et al.*, 2022). Tahapan diawali dengan pengumpulan data dari berbagai sumber hingga menghasilkan suatu informasi dari pengolahan data tersebut. Selanjutnya dilakukan seleksi data dan menghasilkan data yang diinginkan. Tahap selanjutnya yaitu *data preprocessing* yang terdiri dari *data cleaning*, *data integration*, dan *data selection*. Setelah data siap maka data kemudian ditransformasikan ke dalam bentuk yang bisa diolah oleh aplikasi *data mining* sehingga proses *data mining* dapat dijalankan. Hasil dari proses *data mining* yaitu berupa pola yang dapat ditarik kesimpulan dalam suatu interpretasi menjadi suatu informasi

Untuk melakukan pemilihan jumlah *cluster* yang akan digunakan, dapat menggunakan indeks *Davies Bouldin* (DB), dimana pemilihan *cluster* didasarkan pada nilai DB terkecil. Semakin kecil indeks DB, maka hasil pengelompokan yang dilakukan akan semakin baik. Penggunaan indeks DB didasarkan pada perhitungan yang sederhana dan memberikan komputasi *cluster* yang terdistribusi paralel sederhana dan memungkinkan klasifikasi otomatis yang lebih cepat (Rabiah *et al.*, 2015); (Kapoor *et al.*, 2017). Penelitian lain mendefinisikan indeks *Davies Bouldin* sebagai similaritas rata-rata diantara setiap *cluster* dengan *cluster* yang paling mirip, dimana similaritas mengacu pada rasio dari *intra-cluster* dan jarak *inter-cluster*.



Gambar 1. Tahapan proses *data mining*

Persamaan 1 merupakan persamaan yang digunakan untuk menentukan indeks DB dimana  $n$  adalah jumlah *cluster*,  $a_i$  dan  $a_j$  adalah rata-rata jarak antara semua anggota *cluster* dari *cluster*  $i$  dan  $j$ , dan  $d(c_i, c_j)$  adalah jarak rata-rata kedua *centroid* (Vázquez *et al.*, 2020).

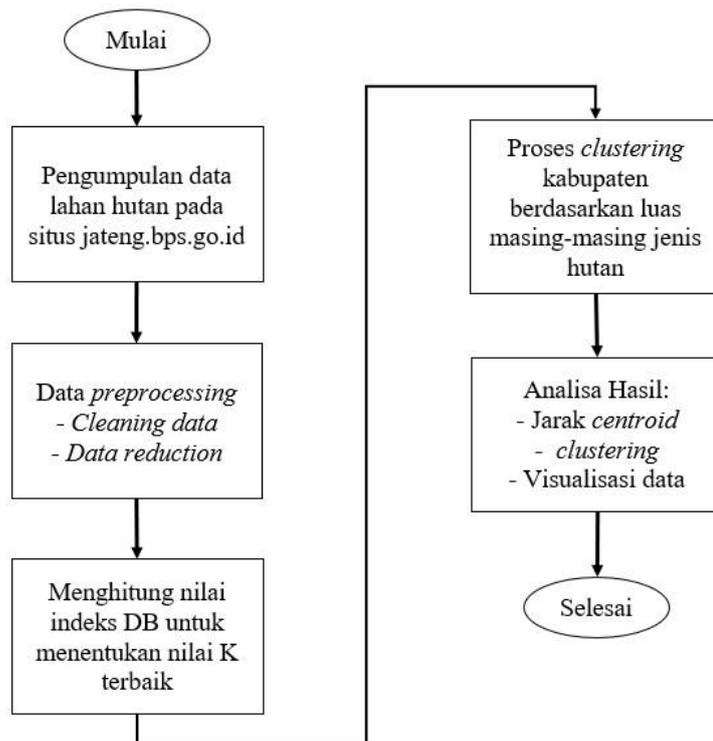
$$DB = \frac{1}{n} \sum_{i=1}^n \max_{i \neq j} \frac{(a_i + a_j)}{d(c_i, c_j)} \dots \dots \dots (1)$$

Gambar 2 merupakan tahapan yang akan dilakukan dalam melakukan penelitian ini. Metodologi pengolahan data yang digunakan adalah metodologi *data mining* yang dimulai dari pengumpulan data luas keempat jenis hutan dari [www.bps.jateng.go.id](http://www.bps.jateng.go.id). Data tersebut selanjutnya melalui *data preprocessing* untuk membersihkan data dari *missing values* dan penggantian nama *attribute* sesuai dengan jenis hutan yang diteliti. Proses pengolahan data dilanjutkan dengan menghitung nilai indeks DB untuk mencari indeks DB untuk *cluster* 2 hingga *cluster* 6 untuk mencari *cluster* terbaik. *Cluster* yang telah dipilih selanjutnya digunakan dalam proses *clustering* untuk mengelompokkan kabupaten berdasarkan luas masing-masing jenis hutan. Hasil yang didapatkan selanjutnya divisualisasikan baik hasil *clustering*, peta wilayah pengelompokan kabupaten, maupun luas masing-

masing hutan di setiap kabupaten. Penentuan jumlah *cluster* yang akan digunakan berdasarkan nilai indeks DB yang didapatkan untuk setiap nilai 'k'. Jumlah *cluster* dengan nilai indeks DB terkecil menjadi pilihan terbaik. Visualisasi yang dilakukan pada penelitian ini menggunakan *Power BI*. Penggunaan *Power BI* didasarkan pada penelitian sebelumnya yang menyatakan keunggulan *Power BI* untuk melakukan proses dan menghasilkan visualisasi data yang lebih baik (Palma-Ruiz *et al.*, 2022).

## HASIL DAN PEMBAHASAN

Dalam melakukan pengolahan data yang digunakan, data yang didapatkan perlu diketahui apakah mengandung *missing data* yang perlu diganti. Pada aplikasi *Rapidminer Studio*, operator *Replace Missing Values* dapat digunakan dan diatur untuk mengganti *missing data* yang ada dengan nilai rata-rata. Selanjutnya dilakukan pemilihan data yang akan digunakan, yaitu data luas hutan antara lain hutan lindung, kawasan perlindungan, kawasan produksi, dan kawasan pengguna lain menggunakan operator *Select Attributes*. Tabel 1 menunjukkan contoh data yang dihasilkan setelah melakukan pemilihan data yang akan digunakan. Data yang dipilih selanjutnya akan dikelompokkan ke dalam kluster-kluster tertentu sesuai dengan nilai data yang digunakan.



Gambar 2. Flowchart penelitian

Tabel 1. Data luas hutan yang telah mengalami tahap *preprocessing* dalam km<sup>2</sup>

| Kabupaten    | Hutan Lindung | Kawasan Perlindungan | Kawasan Produksi | Kawasan Pengguna Lain |
|--------------|---------------|----------------------|------------------|-----------------------|
| Cilacap      | 3655,936      | 21304,08             | 24876,49         | 2508,5                |
| Banyumas     | 7985,22       | 4976,08              | 14928,13         | 800,27                |
| Purbalingga  | 10944,03      | 659                  | 3360,46          | 110,72                |
| Banjarnegara | 2370,93       | 3549,2               | 10941,47         | 233,66                |
| Kebumen      | 3982,56       | 1739,14              | 12654,89         | 583,14                |
| Purworejo    | 3655,936      | 927                  | 7408,03          | 497,64                |
| Wonosobo     | 4255,1        | 3388,5               | 10623,2          | 138,19                |
| Magelang     | 1472,95       | 725,9                | 3023,08          | 53,47                 |
| Boyolali     | 3655,936      | 1690,4               | 11207,28         | 1107,43               |
| Klaten       | 52,3          | 4,6                  | 567,4            | 18,6                  |
| Sukoharjo    | 304,15        | 0,2                  | 70,1             | 990,274               |
| Wonogiri     | 11842,95      | 1009,47              | 7002,13          | 181,91                |
| Karanganyar  | 7080,2        | 4,7                  | 118,7            | 135,6                 |
| Sragen       | 54,5          | 524,4                | 4500,7           | 61,09                 |
| Grobogan     | 3655,936      | 5621,42              | 60249,34         | 3863,1                |
| Bloora       | 3655,936      | 8818,87              | 67404,9          | 5308,81               |
| Rembang      | 2446,46       | 2406,02              | 18215,4          | 948,91                |
| Pati         | 1118,61       | 3153,21              | 16963,57         | 1259,38               |
| Kudus        | 1288,85       | 783,06               | 321,9            | 1241,05               |
| Jepara       | 2707,81       | 5104,15              | 5706,21          | 388,55                |
| Demak        | 3655,936      | 596,42               | 2720,6           | 142                   |
| Semarang     | 1898,87       | 2797,53              | 5336,02          | 947,13                |
| Temanggung   | 3273,32       | 2730,9               | 7263,9           | 238,01                |
| Kendal       | 1702,29       | 3484,3               | 11720,28         | 1719,08               |
| Batang       | 3080,74       | 6691,99              | 6730,58          | 1363,04               |
| Pekalongan   | 1800,87       | 15886,2              | 9959,44          | 358,43                |
| Pemalang     | 5269,42       | 2221,68              | 23104,17         | 2171,62               |
| Tegal        | 2901,52       | 2990,44              | 17245,93         | 670,64                |
| Brebes       | 6252,88       | 6787,05              | 36384,93         | 677,71                |

### Clustering dengan K-Means

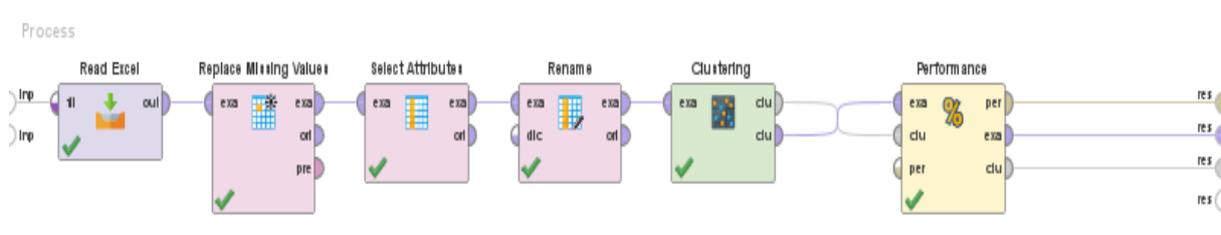
Pada penelitian ini, penentuan jumlah *cluster* yang akan diaplikasikan didasarkan pada nilai DB berdasarkan luas keempat jenis hutan yang diteliti. Tabel 2 memperlihatkan nilai DB untuk *clustering* menggunakan 2 sampai 6 *cluster*. Pada pembagian menjadi 2 *cluster*, nilai DB sebesar 0,436 yang merupakan nilai DB terkecil. Sedangkan untuk pembagian 3 hingga 6 *cluster* masing-masing menghasilkan nilai DB sebesar 0,635, 0,735, 0,593, dan 0,629. Dari nilai DB yang dihasilkan maka dapat disimpulkan bahwa pembagian menjadi 2 *cluster* dengan nilai DB 0,436, menjadi pilihan terbaik dan akan dilanjutkan dengan penentuan jarak dengan titik *centroid* pada masing-masing *attribute* yang diteliti.

Data yang telah siap untuk digunakan selanjutnya akan diolah menggunakan *K-Means*

*clustering*. Gambar 3 menunjukkan model yang digunakan dalam melakukan proses *clustering* menggunakan Rapidminer. Proses diawali dengan memasukkan data menggunakan operator *Read Excel* untuk memasukkan data yang akan diolah dan dilanjutkan dengan pemilihan *attribute*. Data selanjutnya diolah menggunakan operator *Clustering* dan diakhiri dengan operator *Performance*.

Tabel 2. Indeks DB untuk jumlah cluster 2-4

| Jumlah <i>cluster</i> | <i>Davies Bouldin</i> |
|-----------------------|-----------------------|
| 2                     | 0,436                 |
| 3                     | 0,635                 |
| 4                     | 0,735                 |
| 5                     | 0,593                 |
| 6                     | 0,629                 |



Gambar 3. Model analisa *K-Means* menggunakan Rapidminer Studio

Tabel 3. Tabel jarak titik *centroid* antara *cluster* 1 dan *cluster* 2

| <i>Attribute</i>            | <i>Cluster 1</i> | <i>Cluster 2</i> |
|-----------------------------|------------------|------------------|
| Hutan Lindung               | 3556,054         | 4521,584         |
| Kawasan Perlindungan        | 3436,483         | 7075,780         |
| Kawasan untuk Produksi      | 9098,848         | 54679,723        |
| Kawasan untuk Pengguna Lain | 725,705          | 3283,207         |

Tabel 4. Pembagian kabupaten dengan *K-Means*

| <b>Kabupaten</b> | <b>Cluster</b> | <b>Kabupaten</b> | <b>Cluster</b> |
|------------------|----------------|------------------|----------------|
| Cilacap          | 1              | Blora            | 2              |
| Banyumas         | 1              | Rembang          | 1              |
| Purbalingga      | 1              | Pati             | 1              |
| Banjarnegara     | 1              | Kudus            | 1              |
| Kebumen          | 1              | Jepara           | 1              |
| Purworejo        | 1              | Demak            | 1              |
| Wonosobo         | 1              | Semarang         | 1              |
| Magelang         | 1              | Temanggung       | 1              |
| Bojolali         | 1              | Kendal           | 1              |
| Klaten           | 1              | Batang           | 1              |
| Sukoharjo        | 1              | Pekalongan       | 1              |
| Wonogiri         | 1              | Pemalang         | 1              |
| Karanganyar      | 1              | Tegal            | 1              |
| Sragen           | 1              | Brebes           | 2              |
| Grobogan         | 2              |                  |                |

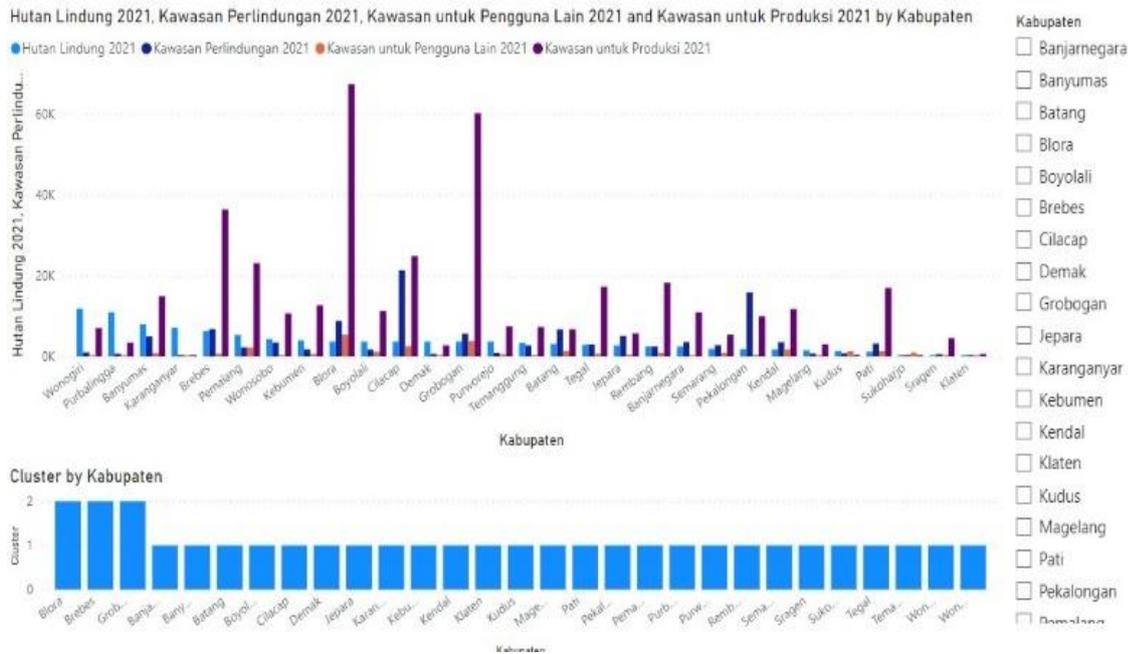
Tabel 3 merupakan tabel jarak antara tiap *attribute* yang terbentuk dengan titik *centroid* pada masing-masing *cluster*. Pada *cluster* 1, jarak nilai rata-rata setiap *attribute* jenis hutan terhadap titik *centroid* adalah 3356,054 pada hutan lindung, 3436,483 pada kawasan perlindungan, 9098,848 pada kawasan untuk produksi, dan 725,705 pada kawasan untuk pengguna lain. Pada *cluster* 2, jarak nilai rata-rata masing-masing *attribute* jenis hutan terhadap titik *centroid* adalah 4521,584 pada hutan lindung, 7075,780 pada kawassan perlindungan, 54679,723 pada kawasan untuk produksi, dan 3283, 207 pada kawasan untuk pengguna lain. Pada tabel ini dapat dilihat bahwa tingkat kedekatan setiap *attribute* pada *cluster* 1 lebih baik dibandingkan pada *cluster* 2.

Tabel 4 menunjukkan pembagian kabupaten ke dalam 2 *cluster* yang ditentukan. Pada tabel tersebut terlihat bahwa hanya ada 3 kabupaten yang masuk ke dalam *cluster* 2, yaitu kabupaten Grobogan, kabupaten Blora, dan kabupaten Brebes, sedangkan sebanyak 26 kabupaten lain yang diteliti dikelompokkan ke dalam cluster 1.

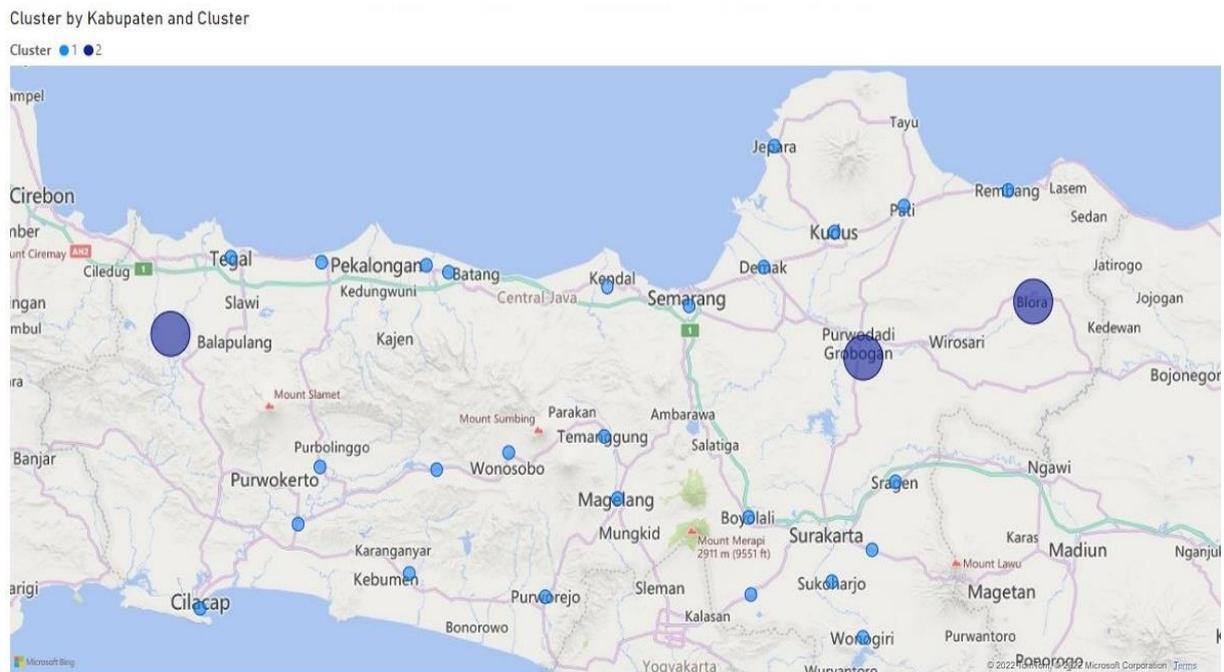
Visualisasi *clustering* yang dilakukan dapat dilihat pada Gambar 4 dan Gambar 5. Gambar 4

merupakan diagram batang yang menunjukkan luas area hutan bedasarkan jenis hutan pada setiap kabupaten dan juga hasil *clustering* yang didapatkan menggunakan aplikasi *Power BI*. Dari gambar yang didapatkan, didapatkan hasil bahwa *cluster* 2 memiliki luas kawasan untuk produksi jauh lebih luas dibandingkan jenis hutan yang lain. Kabupaten pada *cluster* 1 juga memiliki luas hutan produksi yang lebih luas dibandingkan jenis hutan lain, namun perbedaannya tidak sebesar kabupaten pada *cluster* 2.

Gambar 5 memperlihatkan peta pembagian kabupaten berdasarkan *cluster* yang dimiliki menggunakan *Power BI*. Kabupaten pada *cluster* 1 ditandai oleh bulatan kecil berwarna biru, sedangkan pada *cluster* 2, yaitu kabupaten Grobogan, kabupaten Blora, dan kabupaten Brebes ditandai dengan bulatan besar berwarna ungu. Letak geografis yang dimunculkan pada dashboard merupakan gambaran provinsi Jawa Tengah dimana penelitian ini dilakukan, sehingga persebaran kabupaten yang diteliti dapat dilihat secara keseluruhan.



Gambar 4. Diagram luas hutan masing-masing kabupaten dan clustering yang dihasilkan



Gambar 5. Peta pembagian kabupaten berdasarkan clustering yang dilakukan

### KESIMPULAN DAN SARAN

#### Kesimpulan

Pada penelitian ini, clustering dilakukan pada kabupaten berdasarkan luas keempat jenis hutan, yaitu hutan lindung, kawasan perlindungan, kawasan produksi, dan kawasan pengguna lain di Provinsi Jawa Tengah. Pengelompokan dilakukan menjadi 2 cluster berdasarkan nilai indeks DB terkecil sebesar

0,436. Dari hasil perhitungan jarak nilai rata-rata setiap attribute jenis hutan terhadap titik centroid terdekatnya, dihasilkan 26 kabupaten yang termasuk ke dalam cluster 1. Cluster 2 memiliki anggota sebanyak 3 kabupaten, yaitu kabupaten Grobogan, kabupaten Blora, dan kabupaten Brebes. Dari visualisasi yang dihasilkan, dapat disimpulkan bahwa kabupaten Grobogan, Blora, dan Brebes yang merupakan anggota cluster 1 memiliki keunggulan

luas hutan produksi dibandingkan kabupaten lain pada *cluster 2*. Hasil dari *clustering* ini diharapkan menjadi pertimbangan bagi pembuat keputusan seperti pemerintah maupun lembaga-lembaga terkait dalam memaksimalkan pengelolaan hutan berdasarkan keunggulan pada masing-masing *cluster*, dimana akan ada perbedaan kebijakan untuk jenis hutan yang unggul secara luas lahan dibandingkan yang masih perlu ditingkatkan. Sebagai contoh, kabupaten pada *cluster 1* dapat difokuskan pada peningkatan pengelolaan kawasan untuk produksi. Penerapan *clustering* yang dihasilkan pada akhirnya ditampilkan pada *dashboard* diantaranya diagram batang luas hutan untuk setiap kabupaten dan *clustering* yang dihasilkan. Tampilan *dashboard* lain menunjukkan hasil *clustering* menggunakan penanda lingkaran pada letak geografis, dimana untuk kabupaten yang dikelompokkan kedalam *cluster 1* ditandai dengan lingkaran kecil berwarna biru, sedangkan untuk *cluster 2* menggunakan lingkaran yang lebih besar berwarna ungu. Pembuatan visualisasi ini mempermudah pengguna data tersebut dalam melihat pembagian *cluster* yang dibuat.

#### Saran

Hasil penelitian ini diharapkan dapat menjadi salah satu pertimbangan dari pembuat keputusan dalam menentukan kebijakan yang akan diterapkan terkait pengelolaan hutan di wilayah provinsi Jawa Tengah. Sebagai contoh ketersediaan hutan kawasan produksi yang besar pada wilayah kabupaten anggota *cluster 1* dapat dimanfaatkan secara maksimal untuk kesejahteraan masyarakat dengan keputusan-keputusan yang tepat untuk pengelolaan hutan kawasan produksi. Untuk penelitian selanjutnya, diharapkan dapat dilakukan *clustering* menggunakan metode-metode *clustering* yang lain untuk menghasilkan *clustering* yang lebih baik berdasarkan nilai indeks DB yang didapatkan. Pengembangan dalam penelitian ini juga dapat dilakukan dengan melakukan perbandingan hasil *clustering* dengan tahun-tahun sebelumnya atau menambah jumlah daerah yang akan diteliti.

#### DAFTAR PUSTAKA

- Abdul Kadir E, Listia Rosa S, Syukur A, Othman M, Daud H. 2022. Forest fire spreading and carbon concentration identification in tropical region Indonesia. *Alexandria Engineering Journal*, 61(2), 1551–1561. <https://doi.org/10.1016/j.aej.2021.06.064>
- Ahmed M, Seraj R, dan Islam SMS. 2020. The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics (Switzerland)*, 9(8), 1–12. <https://doi.org/10.3390/electronics9081295>
- Alexander C, Korstjens AH, Usher G, Nowak MG, Fredriksson G, Hill RA. 2018. LiDAR patch metrics for object-based clustering of forest types in a tropical rainforest. *International Journal Applied Earth Observation and Geoinformation*, 73:253–261. <https://doi.org/10.1016/j.jag.2018.06.020>
- Badan Pusat Statistik. 2022. *Luas Pembagian Kawasan Hutan Menurut Kabupaten/Kota di Provinsi Jawa Tengah (ha) (Hektar), 2019-2021*. [www.jateng.bps.go.id](http://www.jateng.bps.go.id).
- Hidayat MK dan Fitriana R. 2022. Penerapan Sistem intelijensia bisnis dan *K-Means Clustering* untuk memantau produksi tanaman obat. *Jurnal Teknologi Industri Pertanian*. 32(2):204-219; <https://doi.org/10.24961/j.tek.ind.pert.2022.32.2.204>
- Fitriana R, Saragih J, dan Afrah HB. 2018. Perancangan model sistem intelijensia bisnis untuk menganalisis pemasaran produk roti di pabrik roti menggunakan metode data mining dan cube. *Jurnal Teknologi Industri Pertanian*, 28(1), 113–126. <https://doi.org/10.24961/j.tek.ind.pert.2018.28.1.113>
- Fitriana R, Saragih J, dan Luthfiana N. 2017. Model business intelligence system design of quality products by using data mining in R Bakery Company. *IOP Conference Series: Materials Science and Engineering*, 277(1). <https://doi.org/10.1088/1757-899X/277/1/012005>
- Ida Y, Fujita E, dan Hirose T. 2022. Classification of volcano-seismic events using waveforms in the method of k-means clustering and dynamic time warping. *Journal of Volcanology and Geothermal Research*, 429(January), 107616. <https://doi.org/10.1016/j.jvolgeores.2022.107616>
- Indraputra RA dan Fitriana R. 2020. K-Means clustering data COVID-19. *Jurnal Teknik Industri*. 10 (3):275–282.
- Kapoor S, Zeya I, Singhal C, Nanda SJ. 2017. A Grey wolf optimizer based automatic clustering algorithm for satellite image segmentation. *Procedia Computer Science*. 115: 415–422. <https://doi.org/10.1016/j.procs.2017.09.100>
- Krishnamoorthy M dan Karthikeyan R. 2022. Competence of medicinal plant database using data mining algorithms for large biological databases. *Measurement: Sensors*, 24. 100420. <https://doi.org/10.1016/j.measen.2022.100420>
- Leijten F, Sim S, King H, Verburg PH. 2021. Local deforestation spillovers induced by forest moratoria: Evidence from Indonesia. *Land Use Policy*, 109(September 2020), 105690. <https://doi.org/10.1016/j.landusepol.2021.105690>
- Maier C, Hebermehl W, Grossmann CM, Loft L, Mann C, Hernández-Morcillo M. 2021. Innovations for securing forest ecosystem service provision in Europe – A systematic literature review. *Ecosystem Services*,

- 52(2021).  
<https://doi.org/10.1016/j.ecoser.2021.101374>
- Maryudi A, Laraswati D, Sahide MAK, Giessen L. 2021. Mandatory legality licensing for exports of Indonesian timber products: Balancing the goals of forest governance and timber industries. *Forest Policy and Economics*, 124(January), 102384.  
<https://doi.org/10.1016/j.forpol.2020.102384>
- Muttillainen H dan Vilko J. 2022. Heterogenising forestry value production – Drivers and barriers of entering the non-wood forest products sector. *Current Research in Environmental Sustainability*, 4(November 2021), 100141.  
<https://doi.org/10.1016/j.crsust.2022.100141>
- Naghizadeh A dan Metaxas DN. 2020. Condensed silhouette: An optimized filtering process for cluster selection in K-means. *Procedia Computer Science*, 176, 205–214.  
<https://doi.org/10.1016/j.procs.2020.08.022>
- Nurhayati OD. 2020. Pengolahan Citra untuk Identifikasi Jenis Telur Ayam Lehorn dan Omega-3 Menggunakan k-Means Clustering dan Principle Component Analysis. *Jurnal Sistem Informasi Bisnis*, 01, 84–93.
- Palma-Ruiz JM, Torres-Toukoudis A, González-Moreno SE, Valles-Baca HG. 2022. An overview of the gaming industry across nations: using analytics with power BI to forecast and identify key influencers. *Heliyon*, 8(2).  
<https://doi.org/10.1016/j.heliyon.2022.e08959>
- Pusadan MY, Buliali JL, dan Hari GRV. 2019. Cluster phenomenon to determine anomaly detection of flight route. *Procedia Computer Science*, 161: 516–526.  
<https://doi.org/10.1016/j.procs.2019.11.151>
- Rabaa E, Noura B, dan Adnene C. 2015. Improvements in LEACH based on K-means and Gauss algorithms. *Procedia Computer Science*, 73(Awict), 460–467.  
<https://doi.org/10.1016/j.procs.2015.12.046>
- Rahayu S, Pambudi S, Permadi D, Tata HL, Martini E, Rasnovi S, Nuronia HS, Kindt R, Nugraha M, Dewi S, van Noordwijk M. 2022. Functional trait profiles and diversity of trees regenerating in disturbed tropical forests and agroforests in Indonesia. *Forest Ecosystems*, 9, 100030.  
<https://doi.org/10.1016/j.feecs.2022.100030>
- Rodd M, Fisher M, Monterroso I, Liswanti N, Maryudi A, Larson AM, Mwangi E, Herawati T. 2022. Coordinating forest tenure reform: Objectives, resources and relations in Indonesia, Kenya, Nepal, Peru, and Uganda. *Forest Policy and Economics*, 139(April), 102718.  
<https://doi.org/10.1016/j.forpol.2022.102718>
- Rotich B dan Ojwang D. 2021. Trends and drivers of forest cover change in the Cherangany hills forest ecosystem, western Kenya. *Global Ecology and Conservation*, 30(June), e01755.  
<https://doi.org/10.1016/j.gecco.2021.e01755>
- Santos-Pereira J, Gruenwald L, Bernardino J. 2021. Top data mining tools for the healthcare industry. *Journal of King Saud University - Computer and Information Sciences*, xxx.  
<https://doi.org/10.1016/j.jksuci.2021.06.002>
- Srivastava AK, Jeberson K, dan Jeberson W. 2022. A systematic review on Data Mining Application in Parkinson's disease. *Neuroscience Informatics*. 2(4): 100064.  
<https://doi.org/10.1016/j.neuri.2022.100064>
- Syadzali C, Suryono S, dan Endro SJ. 2020. Business intelligence using the k-nearest neighbor algorithm to analyze customer behavior in online crowdfunding systems. *E3S Web of Conferences*, 202, 1–7.  
<https://doi.org/10.1051/e3sconf/202020216005>
- Ubukata S, Notsu A, dan Honda K. 2021. Objective function-based rough membership C-means clustering. *Information Sciences*. 548: 479–496.  
<https://doi.org/10.1016/j.ins.2020.10.037>
- Vázquez A, López-López N, Sánchez A, Houenou J, Poupon C, Mangin JF, Hernández C, Guevara P. 2020. FFClust: Fast fiber clustering for large tractography datasets for a detailed study of brain connectivity. *NeuroImage*, 220(March).  
<https://doi.org/10.1016/j.neuroimage.2020.117070>
- Weiss G, Hansen E, Ludvig A, Nybakk E, Toppinen A. 2021. Innovation governance in the forest sector: Reviewing concepts, trends and gaps. *Forest Policy and Economics*, 130(May), 102506.  
<https://doi.org/10.1016/j.forpol.2021.102506>