

Penentuan Prioritas Rehabilitasi DAS Menggunakan Algoritma K-Means Clustering

Muhammad Farid Fahmi¹, Yoyon K. Suprpto²

Jurusan Teknik Elektro
Institut Teknologi Surabaya (ITS)
Surabaya, Indonesia

muhammad.farid14@mhs.ee.its.ac.id¹, yoyonsuprpto@ee.its.ac.id²

Abstract—Ketersediaan informasi mengenai tingkat kekritisan lahan yang akurat mempunyai arti yang sangat penting dalam program RHL sehingga Prioritas DAS mana yang akan dilakukan rehabilitasi bisa diketahui. Dari permasalahan diatas diperlukan sebuah cara untuk menentukan prioritas DAS yang direhabilitasi. Adapun metode yang digunakan dalam penelitian ini adalah K-Means Clustering. K-Means Clustering memodelkan dataset menjadi klaster-klaster dimana data pada satu klaster yang memiliki karakteristik sama dan memiliki karakteristik yang berbeda dari klaster lain berdasarkan parameter tingkat kekritisan lahan. Dari penelitian ini diperoleh validasi data dengan penghitungan manual sebesar 80.25% serta didapatkan Kelompok DAS dengan skor rendah untuk semua parameter kekritisan lahan sehingga memiliki tingkat kekritisan lahan tinggi dan menjadi prioritas untuk dilakukan rehabilitasi.

Keywords—Rehabilitasi; DAS; Pengelompokan; K-means

I. PENDAHULUAN

Kerusakan sumber daya hutan berakibat pada menurunnya kemampuan hutan dalam mendukung fungsi ekonomi, sosial dan ekologis. Indikasi kerusakan sumber daya hutan ini dapat dilihat dari menurunnya kualitas Daerah Aliran Sungai (DAS) dan semakin intensnya terjadi bencana alam berupa banjir, kekeringan, dan tanah longsor. Selain itu, kerusakan sumber daya hutan menjadi sorotan dunia internasional sebagai salah satu penyebab perubahan iklim dunia. Dalam rangka mengembalikan kondisi hutan agar mampu berfungsi secara optimal sekaligus untuk mengatasi perubahan iklim, maka diperlukan upaya mitigasi berupa kegiatan Rehabilitasi Hutan dan Lahan (RHL) di 108 DAS prioritas.

Program RHL terlaksana dengan baik apabila informasi obyektif kondisi hutan dan lahan sasaran RHL teridentifikasi secara menyeluruh. Penyediaan data dan informasi sangat diperlukan terutama dalam menunjang formulasi strategi RHL yang berdayaguna dan tepat sasaran, sehingga diharapkan dapat diperoleh acuan dalam pengalokasian sumberdaya secara proporsional.

Ketersediaan informasi mengenai jumlah dan distribusi lahan kritis yang akurat mempunyai arti yang sangat penting dalam memaksimalkan program RHL sehingga Informasi tentang DAS mana yang menjadi prioritas RHL tersedia.

Teknologi data mining adalah salah satu cara mendapatkan informasi yang tersembunyi dari rangkaian data

untuk menjadi pengetahuan yang bisa mendukung sebuah organisasi dalam mengambil keputusan. Salah satu Algoritma dalam data mining adalah *K-Means Clustering*. Dengan menggunakan metode ini, data data yang telah didapatkan dapat dikelompokkan kedalam beberapa *cluster* berdasarkan kemiripan dari data-data tersebut, sehingga data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster* dan yang memiliki karakteristik yang berbeda dikelompokkan dalam *cluster* yang lain yang memiliki karakteristik yang sama [1].

Rumusan permasalahan yang diselesaikan dalam penelitian ini bagaimana menentukan prioritas DAS yang direhabilitasi berdasarkan tingkat kekritisan lahan dengan menggunakan Algoritma *K-means clustering*.

Adapun yang menjadi tujuan dari penelitian ini adalah Untuk melakukan pengelompokan/clustering tingkat kekritisan lahan menggunakan Algoritma *K-means clustering* sehingga diperoleh kelompok DAS yang menjadi prioritas untuk dilakukan rehabilitasi.

II. TINJAUAN PUSTAKA

A. Penelitian Sebelumnya

Penelitian mengenai Data mining dengan algoritma K-means clustering telah banyak dilakukan untuk penyelesaian masalah yang beragam. Salah satunya oleh Oscar [1] yang melakukan clustering untuk menentukan strategi marketing President University. Dari penelitian tersebut didapatkan 3 cluster mahasiswa dengan karakteristik Jurusan, daerah asal dan IPK yang mirip sehingga dari data clustering tersebut bisa dijadikan acuan untuk menentukan strategi promos yang tepat.

Penelitian mengenai Clustering juga dilakukan oleh Agustina [2] yaitu clustering kualitas beras berdasarkan ciri fisik. Hasil dari penelitian ini didapatkan 3 cluster beras dengan kualitas buruk, sedang dan baik.

B. Rehabilitasi DAS (Daerah Aliran Sungai)

Rehabilitasi Hutan dan Lahan (RHL) adalah upaya untuk memulihkan, mempertahankan dan meningkatkan fungsi hutan dan lahan sehingga daya dukung, produktifitas dan peranannya dalam mendukung sistem penyangga kehidupan tetap terjaga. Daerah Aliran Sungai yang selanjutnya disingkat DAS yang tersebar diseluruh wilayah Indonesia, merupakan satu kesatuan

ekosistem alami yang utuh dari ekosistem pegunungan di hulu hingga ekosistem pantai di hilir [3].

Sesuai Peraturan Pemerintah Nomor 37 tahun 2012, Daya Dukung DAS adalah kemampuan DAS untuk mewujudkan kelestarian dan keserasian ekosistem serta meningkatnya kemanfaatan sumber daya alam bagi manusia dan makhluk hidup lainnya secara berkelanjutan. Salah satu sasaran lokasi RHL dipersyaratkan masuk dalam rencana teknik RHL DAS adalah Terdapat hutan rusak dan lahan kritis yang perlu direhabilitasi. Semakin Kritis tingkat kerusakan lahan maka akan menjadi prioritas utama untuk dilakukan Rehabilitasi [3].

C. Parameter Lahan Kritis

Berdasarkan Permenhut Nomor P.32/Menhut-II/2009, Parameter penentu lahan kritis di Hutan Lindung meliputi :

1. Penutupan Lahan

Parameter penutupan lahan dinilai berdasarkan prosentase penutupan tajuk pohon terhadap luas setiap *land system* dan diklasifikasikan menjadi lima kelas. Semakin baik tingkat penutupan lahan maka semakin lahan tersebut tidak mengalami kekritisian dan begitupula sebaliknya. Klasifikasi penutupan lahan untuk masing-masing kelas ditunjukkan pada tabel 1.

Tabel 1. Klasifikasi & Skoring Penutupan Lahan untuk Penentuan Lahan Kritis

Kelas	Prosentase Penutupan Tajuk (%)	Skor
Sangat Baik	> 80	5
Baik	61 - 80	4
Sedang	41 - 60	3
Buruk	21 - 40	2
Sangat Buruk	< 20	1

2. Kemiringan Lereng

Kemiringan lereng adalah perbandingan antara beda tinggi (jarak vertikal) suatu lahan dengan jarak mendatarnya. Besarkemiringan lereng dapat dinyatakan dengan beberapa satuan, diantaranya adalah dengan % (prosen) dan o (derajat). Semakin datar tingkat kemiringan lereng maka semakin lahan tersebut tidak mengalami kekritisian. Klasifikasi kemiringan lereng untuk masing-masing kelas ditunjukkan pada tabel 2.

Tabel 2. Klasifikasi & Skoring Kemiringan Lereng untuk Penentuan Lahan Kritis

Kelas	Kemiringan Lereng (%)	Skor
Datar	< 8	5
Landai	8 - 15	4
Agak Curam	16 - 25	3
Curam	26 - 40	2
Sangat Curam	> 40	1

3. Tingkat Erosi

Tingkat Bahaya Erosi (TBE) dapat dihitung dengan cara membandingkan tingkat erosi di suatu satuan lahan dan kedalaman tanah efektif pada satuan lahan tersebut. Semakin

Semakin Ringan tingkat bahaya erosi maka semakin lahan tersebut tidak mengalami kekritisian. Klasifikasi tingkat bayaha erosi untuk masing-masing kelas ditunjukkan pada tabel 3.

Tabel 3. Klasifikasi & Skoring Tingkat Erosi untuk Penentuan Lahan Kritis

Solum Tanah (cm)	Kelas Erosi				
	I	II	III	IV	V
	Erosi (ton/ha/tahunan)				
	< 15	15 - 60	60 - 180	180 - 480	> 480
Dalam > 90	SR 0	R I	S II	B III	SB IV
Sedang 60 - 90	R I	S II	B III	SB IV	SB IV
Dangkal 30 - 60	S II	B III	SB IV	SB IV	SB IV
Sangat Dangkal <30	B III	SB IV	SB IV	SB IV	SB IV

Ket : 0-SR = Sangat Ringan, I-R = Ringan, II-S = Sedang, III-B= Berat, IV-SB=Sangat Berat

4. Manajemen

Manajemen merupakan salah satu kriteria yang dipergunakan untuk menilai lahan kritis di kawasan hutan lindung, yang dinilai berdasarkan kelengkapan aspek pengelolaan yang meliputi keberadaan tata batas kawasan, pengamanan dan pengawasan serta dilaksanakan atau tidaknya penyuluhan. Semakin baik pengelolaan maka semakin lahan tersebut tidak mengalami kekritisian. Klasifikasi tingkat manajemen untuk masing-masing kelas ditunjukkan pada tabel 4.

Tabel 4. Klasifikasi dan Skoring Manajemen untuk Penentuan Lahan Kritis

Kelas	Besaran / Deskripsi	Skor
Baik	Lengkap *)	5
Sedang	Tidak Lengkap	3
Buruk	Tidak Ada	1

Pada penghitungan manual tingkat kekritisian lahan, dari 4 parameter, bobot masing-masing yaitu Penutupan lahan 50%, kemiringan lereng 20%, tingkat erosi 20% dan manajemen 10%. Nilai dari skor x bobot akan menghasilkan nilai tingkat kekritisian yang ditunjukkan pada table 5.

Tabel 5. Klasifikasi Kekritisian Lahan dikawasan Hutan Lindung.

No	Tingkat Kekritisian Lahan	Besarnya Nilai
1.	Sangat kritis	120 - 180
2.	Kritis	181 - 270
3.	Agak kritis	271 - 360
4.	Potensial kritis	361 - 450
5.	Tidak kritis	451 - 500

D. Data Mining

Data mining adalah metode pengolahan data untuk menemukan pola dan informasi yang tersembunyi dari

sekumpulan data tersebut. Hasil dari pengolahan data dengan metode *data mining* ini dapat menghasilkan informasi yang digunakan untuk mengambil keputusan di masa depan. *Data mining* ini juga dikenal dengan istilah *pattern recognition* [4].

Data mining merupakan metode pengolahan data berskala besar oleh karena itu data mining ini memiliki peranan penting dalam berbagai bidang seperti pemerintahan, industri, keuangan serta ilmu dan teknologi. Secara umum kajian data mining membahas metode-metode seperti prediksi, estimasi, clustering, klasifikasi, asosiasi, regresi, seleksi variable, dan market basket analisis[4].

E. Clustering

Clustering merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*similarity*) antara satu data dengan data yang lain. *Clustering* merupakan salah satu metode *data mining* yang bersifat tanpa arahan (*unsupervised*), maksudnya metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*) serta tidak memerlukan target output. Dalam data mining ada dua jenis metode clustering yang digunakan dalam pengelompokan data, yaitu *hierarchical clustering* dan *non-hierarchical clustering* [4].

Hierarchical clustering adalah suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan kedua. Demikian seterusnya sehingga cluster akan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah *cluster*. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut [5].

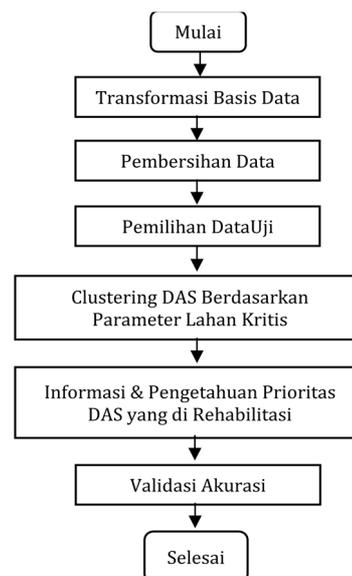
Berbeda dengan metode *hierarchical clustering*, metode *non-hierarchical clustering* justru dimulai dengan menentukan terlebih dahulu jumlah cluster yang diinginkan (dua cluster, tiga cluster, atau lain sebagainya). Setelah jumlah cluster diketahui, baru proses cluster dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan *K-Means Clustering* [5].

III. METODOLOGI

Penelitian ini dilakukan berdasarkan beberapa tahapan yang bertujuan memperlihatkan bagaimana sebuah model *clustering* bisa memberikan informasi tentang karakteristik kelompok tertentu berdasarkan parameter yang ada. Tahapan dalam penelitian ini ditunjukkan pada gambar 1.

A. Transformasi Basis Data

Data yang digunakan dalam penelitian ini adalah file DBF dari Database peta liputan lahan, peta kemiringan lahan, peta erosi dan peta manajemen lahan yang diperoleh dari Balai Pengelolaan Daerah Aliran Sungai (BPDAS) Brantas. Data yang diperoleh masih berupa data dengan banyak atribut sehingga atribut yang tidak digunakan akan dibuang sebagian melalui proses transformasi basis data.



Gambar 1. Diagram alur penelitian

B. Pembersihan Data

Proses selanjutnya adalah melakukan pembersihan data, proses ini bisa menghapus data yang duplikat atau ada field data yang kosong sehingga akan mengurangi jumlah data asal.

C. Pemilihan Data Uji

Proses pemilihan data dilakukan untuk memilih data yang akan diujikan. Mengingat jumlah data DAS yang relatif besar, maka dalam penelitian ini hanya dilakukan clustering terhadap 50 data DAS yang beragam dari 197.346 data DAS yang ada di wilayah kerja BPDAS Brantas.

D. Clustering Data DAS Menggunakan Algoritma K-Means Clustering.

Data yang digunakan dalam penelitian ini adalah file DBF dari Database peta liputan lahan, peta kemiringan lahan, peta erosi dan peta manajemen lahan yang diperoleh dari Balai Pengelolaan Daerah Aliran Sungai (BPDAS) Brantas. Data yang diperoleh masih berupa data dengan banyak atribut sehingga atribut yang tidak digunakan akan dibuang sebagian melalui proses transformasi basis data.

Algoritma *K-Means Clustering* merupakan algoritma pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah K cluster yang sudah ditetapkan di awal. Algoritma *K-Means Clustering* sederhana untuk diimplementasikan dan dijalankan, relatif cepat, mudah beradaptasi serta umum penggunaannya dalam paraktek [6].

Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster*/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan *cluster*/kelompok yang lain sehingga data yang berada dalam satu *cluster*/kelompok memiliki tingkat variasi yang kecil [1].

Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke *cluster* tertentu ditentukan jarak antara data dengan pusat *cluster*. Jarak paling dekat antara satu data dengan satu *cluster* tertentu akan menentukan suatu data masuk dalam *cluster* mana [4]. Perhitungan jarak semua data ke setiap titik pusat *cluster* menggunakan teori jarak Euclidean yang dirumuskan pada persamaan 1.

$$d(x, y) = \sqrt{\sum_{i=1}^r (x_i - y_i)^2} \dots (1)$$

Dimana :

- $d(x, y)$ = Jarak data x ke y
- x_i = Nilai fitur ke- i dari x
- y_i = Nilai fitur ke- i dari y
- r = Jumlah fitur dalam vektor

Suatu data akan menjadi anggota dari *cluster* ke- j apabila jarak data tersebut ke pusat *cluster* ke- j bernilai paling kecil jika dibandingkan dengan jarak ke pusat *cluster* lainnya. Selanjutnya, kelompokkan data-data yang menjadi anggota pada setiap *cluster*[7]. Pusat *cluster* yang baru didapatkan dengan menghitung rata-rata setiap fitur dari semua data yang tergabung dalam setiap *cluster*. Rata-rata sebuah fitur dari semua data dalam *cluster* dinyatakan pada persamaan 2 [6].

$$c_j = \frac{1}{N_k} \sum_{l=1}^{N_k} x_{jl} \quad (2)$$

Dimana :

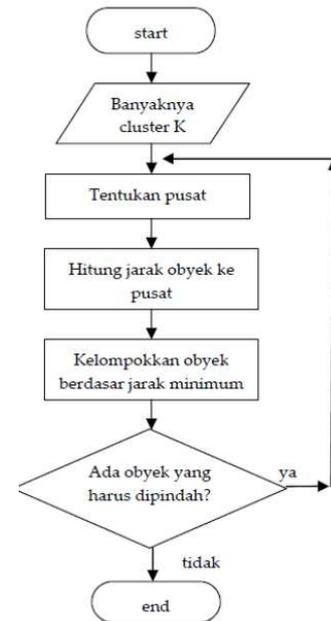
- c_j = Centroid baru
- x_{jl} = Anggota cluster l pada atribut ke- j
- N_k = Jumlah data dalam cluster

Nilai x_{jl} adalah anggota cluster l pada atribut ke- j setelah dihitung dengan persamaan 1. dan N_k adalah jumlah data yang tergabung dalam sebuah cluster setelah dihitung dengan persamaan 1.

Menurut Prasetyo [6], langkah langkah melakukan Algoritma *K-Means Clustering* adalah sebagai berikut :

1. Tentukan nilai k sebagai jumlah cluster.
2. Pilih k dari dataset x sebagai *centroid*.
3. Alokasikan semua data ke *centroid* dengan metrik jarak menggunakan persamaan 1.
4. Hitung kembali *centroid* berdasarkan data yang mengikuti cluster masing-masing menggunakan persamaan 2.
5. Ulangi langkah 3 dan 4 hingga kondisi konvergen tercapai, yaitu (a) perubahan fungsi objektif sudah dibawah ambang batas yang diinginkan atau (b) tidak ada data yang berpindah cluster atau perubahan posisi *centroid* sudah dibawah ambang batas yang ditetapkan.

Diagram alur Algoritma *K-means clustering* ditunjukkan pada gambar 2. [8].



Gambar 2. Flowchart Algoritma K-Means

E. Informasi & Pengetahuan

Informasi yang dihasilkan nantinya akan berupa visualisasi. Visualisasi ini berfungsi untuk mempermudah penulis dalam menganalisis kelompok-kelompok DAS dengan karakteristik yang sama.

F. Validasi Akurasi

Validasi akurasi dilakukan untuk memastikan ketepatan hasil clustering. Dalam penelitian ini validasi akurasi dilakukan menggunakan metode *Mean Squared Error* (MSE) . Rumus dari MSE ditunjukkan pada persamaan 3. [9].

$$MSE = \frac{1}{n} \sum_{t=1}^n (z_t - c_t) \quad (3)$$

Dimana :

- z_t = Data penghitungan manual.
- c_t = Data hasil dihitung dari model
- n = Jumlah data dalam cluster

Nilai z_t diperoleh dari penghitungan manual kekritisan lahan dengan bobot yang sudah ditentukan sedangkan c_t adalah centroid akhir cluster dan n jumlah data dalam cluster yang penghitungannya diperoleh dari persamaan 1.

IV. HASIL DAN PEMBAHASAN

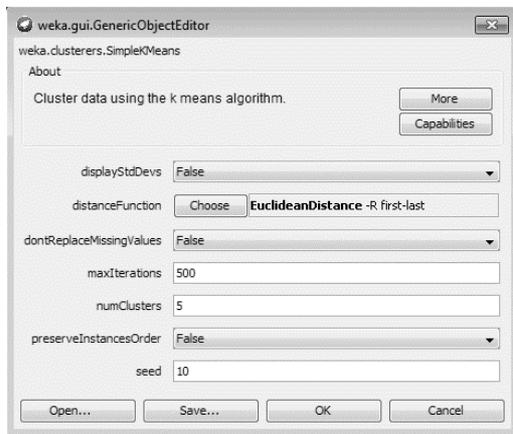
A. Data Uji

Data uji yang digunakan dalam penelitian ini berjumlah 50 record data DAS. Data yang sudah di transformasi dan dibersihkan dari atribut yang tidak dipakai kemudian dipilih secara acak untuk dilakukan clustering. Fitur yang dipakai dalam clustering adalah parameter yang berpengaruh dalam menentukan tingkat kekritisan lahan yaitu skor penutupan lahan, skor lereng, skor erosi dan skor manajemen. data yang sudah dipilih sesuai dengan table 5.

Data Ke-	ID_DAS	SKOR_PL	SKOR_LER	SKOR_ERO	SKOR_MNJ
1	2	1	1	2	3
2	3	1	1	2	3
3	4	1	1	2	3
4	8	1	1	2	3
5	9	1	1	2	3
6	120	5	5	5	5
7	122	5	5	5	5
8	185	3	5	4	5
9	198	3	5	3	5
....
....
50	258	3	3	2	3

B. Clustering Data dengan Algoritma K-Means

Perhitungan pada algoritma k-means melalui bantuan diawali dengan menentukan jumlah cluster atau kelompok data yang dihasilkan nantinya. Dalam penelitian ini menggunakan 5 cluster untuk menentukan tingkat kekritisan lahan. proses penentuan cluster ditunjukkan pada gambar 2.



Gambar 2. Penentuan cluster dan metode yang digunakan untuk menghitung jarak kedekatan

Pada gambar 2. Juga menunjukkan metode apa yang akan digunakan untuk penghitungan jarak. Dalam hal ini kita menggunakan metode Euclidean sesuai dengan persamaan 1.

Setelah kita lakukan proses clustering. Maka didapatkan hasil centroid dan banyaknya data yang masuk untuk masing-masing cluster. Seperti yang ditunjukkan pada gambar 3.

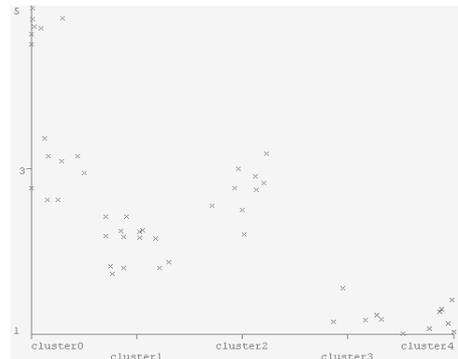
Cluster centroids:	Cluster#				
	0	1	2	3	4
	(15)	(14)	(9)	(5)	(7)
3.9333	2	3	3	1	1
5	3	3	3	1	1
4.0667	2	2	2	2	2
5	3	3	3	3	5

Gambar 3. Centroid dan data untuk masing –masing cluster

Dari Gambar 3. Didapatkan hasil cluster 0 adalah cluster yang memiliki anggota yang paling banyak yaitu 15 data DAS dengan centroid (3.9333;5;4.0667;5).

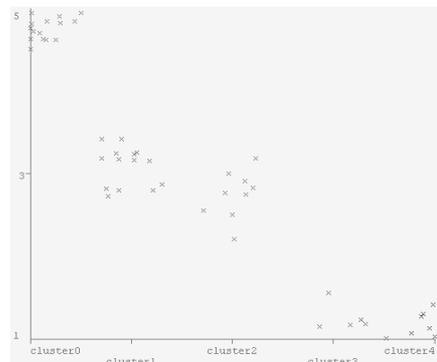
C. Visualisasi Data Cluster

Hasil clustering kemudian divisualisasikan untuk mendapatkan informasi dan mempermudah melakukan analisis. Visualisasi yang di tampilkan dalam bentuk grafik dimana sumbu x adalah data cluster yang terbentuk sedangkan sumbu y adalah skor dari parameter tingkat kekritisan lahan.



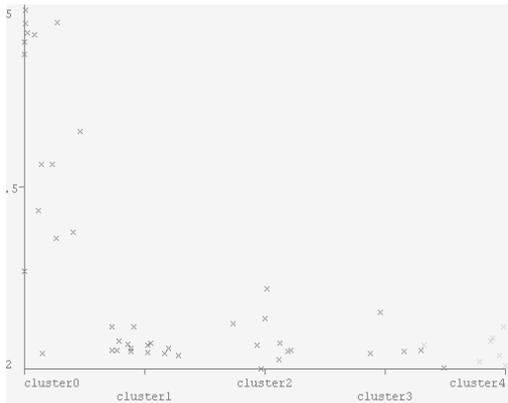
Gambar 4. Cluster berdasarkan Skor Penutupan Lahan

Pada gambar 4. dimana sumbu x adalah data cluster yang terbentuk sedangkan sumbu y adalah skor dari parameter tingkat kekritisan lahan, menunjukkan bahwa cluster 0 adalah kelompok data dengan skor penutupan lahan tinggi dibanding cluster lainnya.



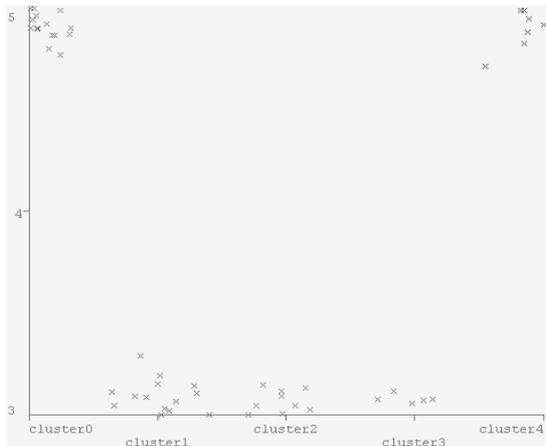
Gambar 5. Cluster berdasarkan Skor Kemiringan Lereng

Pada gambar 5. dimana sumbu x adalah data cluster yang terbentuk sedangkan sumbu y adalah skor dari parameter tingkat kekritisan lahan, Menunjukkan bahwa cluster 0 juga merupakan kelompok data dengan skor kemiringan lereng tinggi dibanding cluster lainnya. Sedangkan untuk skor penutupan lahan terendah di dominasi data di cluster 4 dan cluster 3.



Gambar 6. Cluster berdasarkan Skor Erosi

Pada gambar 6. dimana sumbu x adalah data cluster yang terbentuk sedangkan sumbu y adalah skor dari parameter tingkat kekritisan lahan, menunjukkan bahwa Selain cluster 0, data di cluster lain memiliki skor erosi yang rendah. Artinya DAS di cluster 1, 2, 3 dan 4 memiliki tingkat erosi yang tinggi.



Gambar 7. Cluster berdasarkan Skor Manajemen

Pada gambar 7. dimana sumbu x adalah data cluster yang terbentuk sedangkan sumbu y adalah skor dari parameter tingkat kekritisan lahan, Menunjukkan bahwa selain cluster 0. Cluster 4 adalah kelompok data dengan skor manajemen tinggi, artinya DAS di cluster 0 dan 4 adalah DAS dengan nilai pengelolaan yang baik.

D. Validasi Akurasi

Proses validasi dilakukan dengan menggunakan persamaan 3. Yaitu membandingkan rata-rata pusat cluster/centroid dengan rata-rata data penghitungan manual untuk masing-masing parameter. Hasil validasi dapat dilihat pada table 6.

Tabel 6. Validasi data penghitungan lahan kritis secara manual dengan centroid

Rata-rata /centroid	S_PL	S_LER	S_ERO	S_MNJ	Rata2 validasi
Tidak Kritis	5	5	5	5	80,25%
Cluster 0	3,93	5	4,067	5	
validasi	78,67	100	81,33	100	
Rat-rata	90,00 %				
Potensial Kritis	3	5	5	5	
cluster 1	2	3	2	3	
validasi %	66,66	60	40	60	
Rat-rata	56,66 %				
Agak Kritis	3	3	2	5	
cluster 2	3	3	2	3	
validasi %	100	100	100	60	
Rat-rata	90,00 %				
Kritis	2	3	2	5	
cluster 4	1	1	2	5	
validasi %	50	33,33	100	100	
Rat-rata	70,83 %				
sangat kritis	1	1	2	4	
cluster 3	1	1	2	3	
validasi %	100	100	100	75	
Rat-rata	93,75 %				

V. KESIMPULAN

Algoritma *K-Means Clustering* mampu mengelompokan data DAS menjadi beberapa kelompok sesuai kemiripan dan karakteristik masing-masing dengan Tingkat validasi data antara data hasil dengan hasil penghitungan manual sebesar 80.25 %. Cluster 3 adalah cluster dengan skor terendah untuk semua parameter tingkat kekritisan lahan, ini berarti cluster 3 merupakan kelompok DAS dengan tingkat kekritisan lahan paling tinggi atau sangat kritis sehingga DAS yang ada di kelompok 3 merupakan DAS yang menjadi prioritas untuk dilakukan rehabilitasi.

Saran untuk penelitian berikutnya perlu dicoba clustering DAS dengan metode yang lain untuk membandingkan tingkat akurasi dengan Algoritma *K-Means Clustering*.

REFERENCES

- [1] Oscar, Johan Ong, "Implementasi Algoritma K-Means Clustering untuk menentukan Strategi Marketing President University", Jurnal Ilmiah Teknik Industri, Vol.12 No.1, 2013 .
- [2] Agustina, Silvia "Clustering kualitas beras berdasarkan ciri fisik menggunakan k-means", Program studi ilmu computer Universitas Brawijaya Malang, 2014.
- [3] Dephut, "P.32/Menhut-II/2009 tentang Tata cara penyusunan rencana teknik rehabilitasi hutan dan lahan Daerah Aliran sungai". Jakarta, 2009.
- [4] Santosa B, "Data Mining : Teknik Pemanfaatan Data untuk keperluan Bisnis," Graha Ilmu-Yogyakarta, 2007.
- [5] Santoso, "Statistik Multivariat," Elekmedia Komputindo-Jakarta, 2005.
- [6] Prasetyo, Eko, "Data Mining Mengolah Data menjadi Informasi dengan Matlab," Andi-Yogyakarta, 2009
- [7] Budiman, Irawan, "Data clustering menggunakan metodologi CRISP-DM untuk pengenalan pola proporsi pelaksanaan perguruan tinggi".

Tesis Magister Sistem Informasi Universitas Diponegoro Semarang, 2012.

- [8] Yuhefizar, Budi Santosa, I Ketut Edy P, Yoyon K Suprpto, "Klateralisasi pengunjung web berdasarkan data weblog menggunakan metode *K-means*". Seminar nasional Matematika, 2012.

- [9] Suiang-Shyan LEE, Ja-Chen, LIN, " *An accelerated K-means clustering algorithm using selection and erasure rules* " , Journal of Zhejiang University-Science, 2012.