

BAB III

LANDASAN TEORI

3.1 Tanah Longsor

Tanah longsor adalah perpindahan material pembentuk lereng berupa batuan, bahan rombakan, tanah atau material laporan bergerak ke bawah atau keluar lereng. Secara geologis tanah longsor adalah suatu peristiwa geologi dimana terjadi pergerakan tanah seperti jatuhnya bebatuan atau gumpalan besar tanah (Nandi, 2007:6). Menurut Karnawati (2005), gerakan masa tanah terjadi akibat terganggunya kestabilan tanah atau batuan penyusun lereng, sehingga masa tanah atau bahan penyusun lereng maupun percampuran keduanya mengalami gerakan menuruni lereng.

Pada prinsipnya tanah longsor terjadi bila gaya pendorong pada lereng lebih besar dari pada gaya penahan. Gaya penahan pada umumnya dipengaruhi oleh kekuatan batuan dan kepadatan tanah. Sedangkan gaya pendorong dipengaruhi oleh besarnya sudut lereng, beban serta berat jenis batuan. Selain itu, tanah longsor juga disebabkan karena terjadinya getaran. Getaran yang terjadi biasanya diakibatkan oleh getaran gempa bumi.

Proses terjadinya tanah longsor dapat di jelaskan sebagai berikut, air yang meresap ke dalam tanah akan menambah bobot tanah. Jika air tersebut menembus sampai tanah kedap air yang berperan sebagai bidang gelincir, maka tanah menjadi licin dan tanah pelapukan di atasnya akan bergerak mengikuti lereng dan luar lereng.

3.2 Penyebab Terjadinya Tanah Longsor

Menurut Nandi (2007:6) gejala umum tanah longsor ditandai dengan munculnya retakan-retakan di lereng yang sejajar dengan arah tebing, biasanya terjadi setelah hujan, munculnya mata air baru secara tiba-tiba dan tebing rapuh serta kerikil mulai berjatuhan. Faktor lainnya adalah sebagai berikut :

1. Hujan

Musim kering yang panjang akan menyebabkan terjadinya penguapan air di permukaan tanah dalam jumlah besar. Hal ini mengakibatkan munculnya pori-pori tanah hingga terjadi retakan dan merekahnya tanah ke permukaan. Ketika hujan, air akan menyusup kebagian yang retak sehingga tanah dengan cepat mengembang kembali.

Pada awal musim hujan, intensitas hujan yang tinggi biasanya sering terjadi, sehingga kandungan air pada tanah menjadi jenuh dalam waktu yang singkat. Hujan lebat pada awal musim dapat menimbulkan longsor karena melalui tanah yang merekah air akan masuk dan terakumulasi dibagian dasar lereng, sehingga menimbulkan gerakan lateral.

2. Lereng Terjal

Lereng atau tebing yang terjal akan memperbesar gaya pendorong. Lereng yang terjal terbentuk karena pengikisan air sungai, mata air, air laut, dan angin. Kebanyakan sudut lereng yang menyebabkan longsor adalah 180° apabila ujung lerengnya terjal dan bidang longsornya datar.

3. Tanah yang Kurang Padat dan Tebal

Jenis tanah yang kurang padat adalah tanah lempung atau tanah liat dengan ketebalan lebih dari 2,5 m dari sudut lereng lebih dari 220° . Tanah jenis ini memiliki potensi untuk terjadinya tanah longsor terutama bila terjadi hujan. Selain itu tanah ini sangat rentan terhadap pergerakan tanah karena menjadi lembek terkena air dan pecah ketika hawa terlalu panas.

4. Batuan yang Kurang Kuat

Batuan endapan gunung api dan sedimen berukuran pasir dan campuran antara kerikil, pasir dan lempung umumnya kurang kuat. Batuan tersebut akan mudah menjadi tanah apabila mengalami proses pelapukan dan umumnya rentan terhadap tanah longsor bila terdapat pada lereng yang terjal.

5. Jenis Tata Lahan

Tanah longsor banyak terjadi di daerah lahan persawahan, perladangan dan adanya genangan air di lereng yang terjal. Pada lahan persawahan akarnya kurang kuat untuk mengikat butir tanah yang membuat tanah menjadi lembek dan jenuh

dengan air sehingga mudah longsor. Sedangkan untuk daerah perladangan penyebabnya adalah karena akar pohonnya tidak dapat menembus bidang longsor yang dalam dan umumnya terjadi di daerah longsor lama.

6. Getaran

Getaran yang terjadi biasanya diakibatkan oleh gempa bumi, ledakan, getaran mesin dan getaran lalu lintas kendaraan. Akibat yang ditimbulkan adalah tanah, badan jalan, lantai dan dinding rumah menjadi retak.

3.3 Jenis-jenis Tanah Longsor

Menurut (Cruden dan Varnes 1992, dalam Hary, 2006:15) tanah longsor dikelompokkan menjadi jatuhan, robohan, longsor, sebaran dan aliran. Masing masing tipe terjadi pada medan dengan karakteristik yang berbeda satu dengan yang lainnya, hal ini karena bencana gerakan tanah disebabkan oleh banyak faktor.

1. Jatuhan (*falls*)

Jatuhan (*falls*) adalah gerakan jatuh material pembentuk lereng (tanah atau batuan) di udara dengan atau tanpa adanya interaksi antara bagian-bagian material yang longsor. Jatuhan terjadi tanpa adanya bidang longsor, dan banyak terjadi pada lereng terjal atau tegak yang terdiri dari batuan yang mempunyai bidang-bidang tidak menerus (diskontinuitas). Jatuhan pada tanah biasanya terjadi bila material mudah tererosi terletak di atas tanah yang lebih tahan erosi, contohnya jika lapisan pasir bersih atau lanau berada di atas lapisan lempung *overconsolidated* (Bazett et al, 1961; Skempton dan La Rochelle 1965).

Jatuhan batuan dapat terjadi pada semua jenis batuan dan umumnya terjadi akibat pelapukan, perubahan temperatur, tekanan air atau penggalan/penggerusan bagian bawah lereng. Jatuhan terjadi di sepanjang kekar, bidang dasar, atau zona patahan lokal.

2. Robohan (*topples*)

Robohan (*topples*) adalah gerakan material robohan dan biasanya terjadi pada lereng batuan yang sangat terjal sampai tegak yang mempunyai bidangbidang ketidakterusan yang relatif vertikal. Tipe gerakan hampir sama

dengan jatuhan, hanya gerakan batuan longsor adalah mengguling hingga roboh, yang berakibat batuan lepas dari permukaan lerengnya. Faktor utama yang menyebabkan robohan, adalah seperti halnya kejadian jatuhan batuan, yaitu yang mengisi retakan.

3. Longsoran (*slides*)

Longsoran (*slides*) adalah gerakan material pembentuk lereng yang diakibatkan oleh terjadinya kegagalan geser, di sepanjang satu atau lebih bidang longsor. Massa tanah yang bergerak bisa menyatu atau terpecah-pecah. Perpindahan material total sebelum longsoran bergantung pada besarnya regangan untuk mencapai kuat geser puncaknya dan pada tebal zona longsornya. Perpindahan total lebih kecil pada lempung *normally consolidated* daripada lempung kaku *overconsolidated*.

Berdasarkan geometri bidang gelincirannya, longsoran dibedakan dalam dua jenis yaitu: (Hary, 2006:21)

a. Longsoran dengan bidang longsor lengkung atau longsoran rotasional (*rotational slides*).

Longsoran rotasional (*rotational slides*) mempunyai bidang longsor melengkung ke atas, dan sering terjadi pada massa tanah yang bergerak dalam satu kesatuan. longsoran rotasional murni (*slump*) terjadi pada material yang relatif homogen seperti timbunan buatan (tanggul).

b. Longsoran dengan bidang gelincir dasar atau longsoran translasional (*translational slides*).

Longsoran translasional merupakan gerakan di sepanjang diskontinuitas atau bidang lemah yang secara pendekatan sejajar dengan permukaan lereng, sehingga gerakan tanah secara translasi. Dalam tanah lempung, translasi terjadi di sepanjang lapisan tipis pasir atau lanau, khususnya bila bidang lemah tersebut sejajar dengan lereng yang ada. Longsoran translasi lempung yang mengandung lapisan pasir atau lanau, dapat disebabkan oleh tekanan air pori yang tinggi dalam pasir atau lanau tersebut.

4. Sebaran (*spreads*)

Sebaran yang termasuk longsoran translasional juga disebut sebaran lateral (*lateral spreading*), adalah kombinasi dari meluasnya massa tanah dan turunnya massa batuan terpecah-pecah ke dalam material lunak di bawahnya (Cruden dan Varnes, 1992 dalam Hary, 2006). Permukaan bidang longsor tidak berada di lokasi terjadinya geseran terkuat.

5. Aliran (*flows*)

Aliran (*flows*) adalah gerakan hancuran material ke bawah lereng dan mengalir seperti cairan kental. Aliran sering terjadi dalam bidang geser relatif sempit. Material yang terbawa oleh aliran dapat terdiri dari berbagai macam pertikel tanah (termasuk batu-batu besar), kayu-kayu, ranting dan lain-lain.

3.4 Dampak Tanah Longsor

Menurut Nandi (2007:17) banyak dampak yang ditimbulkan akibat terjadinya tanah longsor baik dampak terhadap kehidupan manusia, hewan dan tumbuhan maupun dampak terhadap keseimbangan lingkungan.

1. Dampak Terhadap Kehidupan

Terjadinya bencana tanah longsor memiliki dampak yang sangat besar terhadap kehidupan, khususnya manusia. Bila tanah longsor itu terjadi pada wilayah yang memiliki kepadatan penduduk yang tinggi, maka korban jiwa yang ditimbulkan akan sangat besar, terutama bencana tanah longsor itu terjadi secara tiba-tiba tanpa diawali adanya tanda-tanda akan terjadinya tanah longsor.

Adapun dampak yang ditimbulkan dengan terjadinya tanah longsor terhadap kehidupan adalah sebagai berikut.

- a. Bencana longsor banyak menelan korban jiwa.
- b. Terjadinya kerusakan infrastruktur publik seperti jalan, jembatan dan sebagainya.
- c. Kerusakan bangunan seperti gedung perkantoran dan perumahan penduduk serta sarana peribadatan.
- d. Menghambat proses aktivitas manusia dan merugikan baik masyarakat yang terdapat di sekitar bencana maupun pemerintahan.

2. Dampak Terhadap Lingkungan

Adapun dampak yang ditimbulkan terhadap lingkungan akibat terjadinya tanah longsor adalah sebagai berikut:

- a. Terjadinya kerusakan lahan.
- b. Hilangnya vegetasi penutup lahan.
- c. Terganggunya keseimbangan ekosistem.
- d. Lahan menjadi kritis sehingga cadangan air bawah tanah menipis.
- e. Terjadinya tanah longsor dapat menutup lahan yang lain seperti sawah, kebun dan lahan produktif lainnya.

3.5 Statistika Deskriptif

Statistika deskriptif adalah statistik yang berfungsi untuk mendeskripsikan atau memberi gambaran terhadap obyek yang diteliti melalui data sampel atau populasi sebagaimana adanya, tanpa melakukan analisis dan membuat kesimpulan yang berlaku untuk umum (Sugiyono, 2010). Sedangkan menurut Wahyuni (2011), statistika deskriptif merupakan statistik yang fokus perhatiannya dan penganalisisan data saja, tanpa berusaha untuk menarik kesimpulan terhadap populasinya. Data tersebut biasanya disajikan dengan menggunakan tabel atau grafik.

1. Tabel

Tabel adalah salah satu cara penyajian data, yang mana terdapat kolom dan baris didalamnya serta terdapat judul tabel (Sugiyono, 2010).

2. Diagram Pareto

Diagram pareto adalah serangkaian diagram batang yang menggambarkan frekuensi atau pengaruh dari proses atau keadaan atau masalah. Diagram diatur mulai dari yang paling tinggi sampai paling rendah dari kiri ke kanan atau atas ke bawah (Shahindra, 2008).

3. *Mean*

Mean merupakan teknik penjelasan kelompok yang didasarkan atas nilai rata-rata dari kelompok tersebut. Rata-rata (*mean*) ini didapat dengan

menjumlahkan data seluruh individu dalam kelompok itu, kemudian dibagi dengan jumlah individu yang ada pada kelompok tersebut (Sugiyono, 2010).

$$\text{Mean} = \frac{\sum_{i=1}^n X_i}{n} \quad \dots 3.1$$

dengan,

X_i = individu x ke i sampai ke n

n = banyaknya individu

3.6 Data Mining

Data mining adalah suatu proses dalam menemukan korelasi, pola dan tren baru pada suatu data dengan cara memilah-milah sejumlah data besar yang disimpan dalam sebuah tempat penyimpanan data dengan menggunakan suatu teknologi seperti statistika dan matematika (Larose, 2005).

Sumber data dalam *data mining* dapat berupa *database*, data *warehouse*, *website*, dan data lainnya. Data yang paling sering digunakan adalah *database*, data *warehouse*, data transaksi, data spasial, data *stream*, data *sequence*, grafik atau data *network* (Han dan Kamber, 2012). Beberapa teknik *data mining* yang umum digunakan menurut Larose (2005) antara lain:

1. Deskripsi digunakan untuk mencari pola dan kecenderungan yang terdapat dalam data.
2. Estimasi hampir sama dengan klasifikasi, namun variabel target dalam estimasi lebih ke arah numerik daripada kategorik.
3. Prediksi juga hampir sama dengan estimasi dan klasifikasi, namun dalam prediksi nilai dari hasil akan ada di masa mendatang.
4. Klasifikasi adalah fungsi pembelajaran yang memetakan (mengklasifikasi) sebuah unsur (*item*) data ke dalam salah satu dari beberapa kelas yang sudah didefinisikan.
5. Pengelompokkan atau *clustering* merupakan tugas deskripsi yang banyak digunakan dalam mengidentifikasi sebuah hubungan terbatas pada kategori atau *cluster* untuk mendeskripsikan data yang ditelaah.

6. Asosiasi dalam data mining bertugas untuk menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

3.7 Clustering

3.7.1 Pengenalan Clustering

Clustering adalah pengelompokan dari *record*, observasi-observasi atau kasus-kasus ke kelas yang memiliki kemiripan pada objek-objeknya. *Cluster* adalah koleksi dari *record* yang mirip, dan tidak mirip dengan *record* dari *cluster* lain. *Clustering* berbeda dengan klasifikasi, dalam hal tidak ada variabel target untuk *clustering*. *Clustering* tidak mengklasifikasikan, meramalkan, atau memprediksi nilai dari sebuah variabel target. Algoritma-algoritma *clustering* digunakan untuk menentukan segmen keseluruhan himpunan data menjadi sub *group* yang relatif sama atau *cluster* dengan kesamaan *record* dalam *cluster* dimaksimumkan dan kesamaan *record* di luar *cluster* diminimumkan (Larose, 2005).

Tujuan *clustering* adalah mengelompokkan obyek atas dasar karakteristik yang dimiliki. Analisis *cluster* mengelompokkan obyek (responden, produk, atau entitas lainnya) sehingga masing-masing obyek mempunyai kemiripan dengan yang lain dalam suatu *cluster*. Hasil *cluster* suatu obyek harus memiliki internal (*within cluster*) homogenitas yang tinggi dan memiliki eksternal (*between cluster*) heterogenitas yang tinggi. Jika pengelompokan berhasil, maka obyek dalam satu *cluster* akan saling dekat satu sama lain jika diplot secara geometri dan *cluster* yang berbeda akan saling menjauh satu sama lain (Ghozali, 2005 dalam Muthiah, 2013).

Konsep dasar pengukuran analisis *cluster* adalah konsep pengukuran jarak (*distance*) dan kesamaan (*similarity*). Konsep pengukuran jarak adalah ukuran tentang jarak pisah antar obyek, sedangkan konsep pengukuran kesamaan adalah berbicara tentang suatu ukuran kedekatan. Konsep ini penting karena pengelompokan pada analisis *cluster* didasarkan pada kedekatan. Pengukuran jarak (*distance type measure*) digunakan untuk data-data yang bersifat matriks,

sedangkan pengukuran kesesuaian (*matching type measure*) digunakan untuk data-data yang bersifat kualitatif (Fauzy, 1999 dalam Muthiah, 2013).

Menurut Supranto (2004), konsep pengukuran jarak merupakan pendekatan yang paling biasa dalam mengukur kemiripan antara pasangan objek. Pasangan objek dengan jarak yang lebih pendek akan lebih mirip dibandingkan pasangan objek yang jaraknya lebih panjang. Ada beberapa cara untuk mengukur jarak antara dua objek, diantaranya:

1. Jarak *euclidean* adalah jumlah kuadrat perbedaan atau deviasi di dalam nilai untuk setiap variabel (Supranto, 2004). Menurut Supianto (2014), jarak *euclidean* merupakan pengukuran jarak yang paling umum digunakan pada data numerik.

$$d_{(ij)} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2}$$

...3.2

dengan,

$d_{(ij)}$ = jarak antara objek i dan objek j

x_{ik} = nilai objek i pada variabel ke k

x_{jk} = nilai objek j pada variabel ke k

n = banyak variabel yang diamati

Contoh:

Tabel 3.1 Contoh Data Perhitungan *Euclidean*

Daerah	X_1	X_2
A	6	6
B	8	10
C	6	4
D	4	6

Perhitungan jarak :

$$d_{A,B} = (6-8)^2 + (6-10)^2 = \sqrt{20} = 4,472$$

Perhitungan jarak :

Dengan cara yang sama dapat dilakukan dalam perhitungan objek-objek lainnya. Jarak antar objek dapat dilihat secara keseluruhan pada tabel 3.2 berikut:

Tabel 3.2 Matriks Jarak *Euclidean*

	A	B	C	D
A	0	20	4	4
B	20	0	40	32
C	4	40	0	8
D	4	32	8	0

2. *Manhattan distance* atau *the city blok distance* adalah jumlah perbedaan mutlak/absolut di dalam nilai untuk setiap variabel (Supranto, 2004).

$$d_{(ij)} = \sum_{k=1}^n |x_{ik} - x_{jk}| \quad \dots 3.3$$

dengan,

$d_{(ij)}$ = jarak antara objek i dan objek j

x_{ik} = nilai objek i pada variabel ke k

x_{jk} = nilai objek j pada variabel ke k

n = banyak variabel yang diamati

Contoh:

$$d_{A,B} = |6-8| + |6-10| = 6$$

Dengan cara perhitungan yang sama, jarak antar objek dapat dilihat secara keseluruhan pada tabel 3.3 berikut:

Tabel 3.3 Matriks Jarak *Manhattan*

	A	B	C	D
A	0	6	2	2
B	6	0	8	8
C	2	8	0	4
D	2	8	4	0

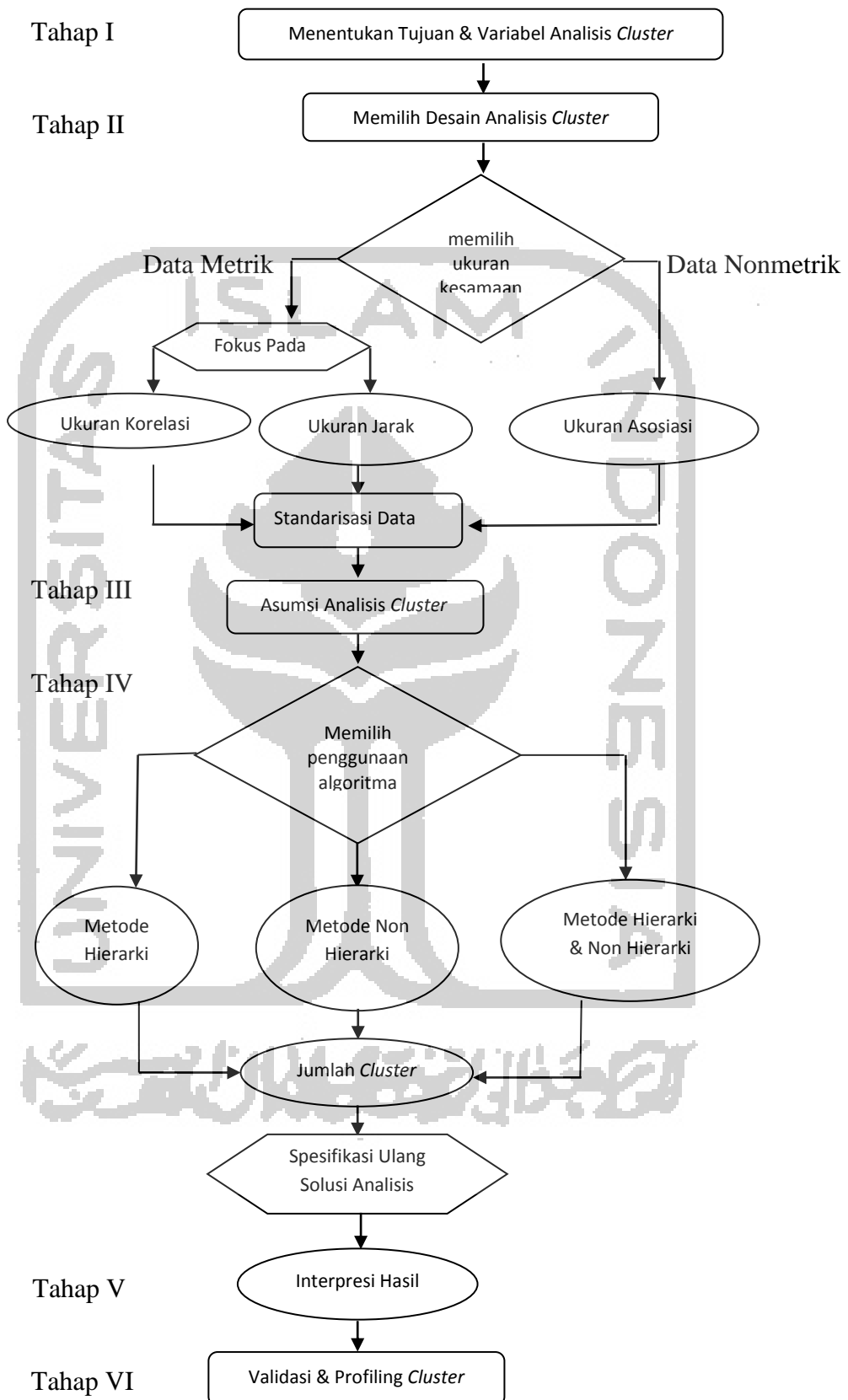
3.7.2 Metode-metode dalam *Clustering*

Secara umum metode utama *clustering* dapat di klasifikasikan menjadi kategori-kategori berikut (Pratama, 2015):

1. Metode partisi adalah sebuah n objek yang dipartisi menjadi sejumlah k *cluster* ($k \leq n$), dimana partisi yang terbentuk harus memenuhi syarat yaitu setiap *cluster* harus berisi minimal satu objek dan setiap objek harus termasuk tepat satu *cluster*.
2. Metode hirarki yaitu membuat sebuah dekomposisi berhirarki dari himpunan data atau objek menggunakan beberapa kriteria. Metode ini memiliki dua jenis pendekatan yaitu *agglomerative* (metode penggabungan) dan *divisive* (metode pembagian). *Agglomerative* dimulai dengan titik-titik sebagai *cluster* individu. Pada setiap tahap dilakukan penggabungan setiap pasangan titik pada *cluster* sampai hanya satu titik atau *cluster* yang tertinggal. Sedangkan *divisive* dimulai dengan satu *cluster* besar yang berisi semua titik data yang kemudian dibagi menjadi dua *cluster*, dan seterusnya. Pada setiap langkah dilakukan pemecahan sebuah *cluster* sampai setiap *cluster* berisi sebuah titik.
3. Metode berdasarkan kepekatan merupakan metode yang berdasarkan pada konektivitas dan fungsi kepadatan.
4. Metode berdasarkan *grid* merupakan pendekatan yang berdasarkan pada struktur *multiple-level granularity*.
5. Metode berdasarkan model yaitu sebuah model yang ide dasarnya menemukan model yang cocok untuk tiap *cluster*.

3.7.3 Proses Pengambilan Keputusan Analisis *Cluster*

Untuk memudahkan dalam pengambilan keputusan dalam analisis *cluster*, berikut ditampilkan bagan pengambilan keputusannya.



Gambar 3.2 Proses Pengambilan Keputusan Analisis *Cluster*

(Sumber: *Multivariate Analysis, Hair dkk 2006 dalam Nurullina 2010*)

3.8 Self Organizing Maps (SOM)

Jaringan Kohonen *SOM* diperkenalkan pada tahun 1982 oleh peneliti Finlandia yaitu Tuevo Kohonen. Meskipun awalnya diterapkan untuk analisis gambar dan analisis suara, namun jaringan Kohonen merupakan mekanisme yang efektif untuk analisis *clustering* (Larose, 2005). Kohonen *SOM* merupakan suatu jaringan yang tidak membutuhkan suatu pengawasan khusus, karenanya diberi nama *self organizing*. Kata *maps* berarti bahwa metode ini menggunakan *map* dalam pembobotan input data. Tiap *node* dalam jaringan *SOM* berusaha untuk menjadi seperti input yang telah diberikan pada jaringan tersebut. *SOM* juga biasa disebut dengan *Self Organizing Feature Maps*, maksudnya adalah bahwa *SOM* menggunakan prinsip "*feature*" atau ciri khusus dalam prinsip dasarnya yang membuatnya berbeda dibandingkan metode yang lain (Guthikonda, 2005).

Metode kohonen memungkinkan untuk menggambarkan data multidimensi ke dalam dimensi yang lebih kecil, biasanya satu atau dua dimensi serta memiliki kemampuan untuk belajar secara mandiri (*unsupervised learning*). Artinya, sebuah jaringan akan belajar dengan dibekali pengetahuan dasar (parameter-parameter jaringan) tanpa adanya pengetahuan awal lebih dulu mengenai segmen dan karakteristiknya serta tanpa harus mengetahui berapa kelompok yang akan dibentuk, dan kemudian mengorganisasikan sendiri hubungan-hubungan interkoneksi dalam dirinya atas masukan yang diberikan sehingga dengan demikian target tidak dibutuhkan (Septianusa dan Supriyaningsih, 2014).

Bullinaria (2004) menyebutkan bahwa, salah satu komponen utama yang penting dari *unsupervised system* adalah *competitive learning* yaitu *output* yang berupa neuron akan bersaing dengan neuron yang lain untuk diaktifkan. Kemudian nanti hanya terdapat satu neuron yang diaktifkan dalam satu waktu. Neuron aktif ini disebut sebagai *winning neuron*.

Seperti telah disebutkan sebelumnya bahwa *SOM* menggunakan pola *unsupervised learning* yang mana salah satu komponen utama dari pola tersebut adalah *competitive learning*. Menurut Haykin (1999), terdapat empat komponen penting dalam *SOM* yaitu:

1. *Initialization*. Menentukan bobot acak awal secara *random*.
2. *Competition*. Pada tahap ini untuk setiap pola input, neuron menghitung nilai masing-masing fungsi diskriminan yang memberi dasar untuk kompetisi. Neuron tertentu dengan nilai terkecil dari fungsi diskriminan dinyatakan sebagai pemenang (*winning neuron*).
3. *Cooperation*. *Winning neuron* menentukan lokasi spasial dari lingkungan topologi *excited neuron* untuk memberi dasar kerjasama dalam suatu lingkungan neuron.
4. *Adaption: Excited neuron* menurunkan nilai fungsi diskriminan yang berkaitan dengan pola input melalui penyesuaian bobot terkait sehingga respon dari *winning neuron* ke aplikasi berikutnya dengan pola input yang sama akan meningkat.

SOM digunakan sebagai visualisasi, alat analisis untuk data berdimensi tinggi, *clustering*, *classification*, *dimensionality reduction*, *sampling*, *vector quantization* dan *data mining*. (Bacao dan Lobo). *SOM* dianggap sebagai bentuk spasial dari analisis kelompok *K-Means*. Analoginya, setiap unit sesuai dengan sebuah *cluster* dan jumlah *cluster* ditentukan oleh ukuran grid yang biasanya diatur dalam bentuk persegi atau heksagonal. Grid ini digunakan dalam proses pemetaan. Jadi ketika objek dua dimensi sangat mirip, maka posisinya dalam pemetaan akan sangat berdekatan (Wehrens dan Buydens, 2007). Keuntungan utama dari *SOM* dibandingkan *K-Means* adalah cenderung lebih sedikit mendapatkan hasil percabangan daripada menggunakan algoritma *K-Means*, dan dapat digunakan sebagai algoritma inisialisasi yang baik untuk metode *K-Means*. Kelebihan lainnya adalah diperolehnya suatu urutan topologis yang biasanya *cluster* yang mirip disusun bersama (Lobo, 2009).

3.8.1 Tahapan dalam SOM

Adapun tahapan dalam pola jaringan Kohonen SOM, menurut Siang (2009) dalam Pratama (2015) adalah dengan melakukan langkah berikut.

1. Inisialisasi berupa bobot (w_{ij}) yang diperoleh secara acak untuk tiap *node*. Setelah bobot (w_{ij}) diberikan maka jaringan diberikan *input* (x_i).

Contoh:

Tabel 3.4 Contoh bobot acak

Bobot acak	X_1	X_2	X_3	X_4
1	0,3	0,5	0,4	0,3
2	0,2	0,3	0,6	0,2
3	0,4	0,5	0,3	0,4
4	0,1	0,5	0,8	0,4

2. Setelah *input* diterima jaringan akan melakukan perhitungan jarak *vector* $d_{(j)}$ yang didapat dengan menjumlahkan selisih antara vektor bobot (w_{ij}) dengan vektor *input* (x_i).

$$d_j = \sum_I (w_{ij} - x_i)^2 \quad \dots 3.4$$

Contoh:

Tabel 3.5 Contoh Data Untuk Analisis SOM

Provinsi	X_1	X_2	X_3	X_4
Aceh	-0,22844777	-0,25024328	-0,22180905	0,3492386

$$\begin{aligned}
 d_j &= \sum_I (w_{ij} - x_i)^2 \\
 &= (0,3 - (-0,228))^2 + (0,5 - (-0,25))^2 + (0,4 - (-0,222))^2 + (0,3 - 0,35)^2 \\
 &= 1,231 \text{ (bobot acak 1)} \\
 d_j &= (0,2 - (-0,228))^2 + (0,3 - (-0,25))^2 + (0,6 - (-0,222))^2 + (0,2 - 0,35)^2 \\
 &= 1,284 \text{ (bobot acak 2)} \\
 d_j &= (0,4 - (-0,228))^2 + (0,5 - (-0,25))^2 + (0,3 - (-0,222))^2 + (0,4 - 0,35)^2 \\
 &= 1,232 \text{ (bobot acak 3)}
 \end{aligned}$$

$$d_j = (0,1-(-0,228))^2 + (0,5-(-0,25))^2 + (0,8-(-0,222))^2 + (0,4-0,35)^2$$

$$= 1,717 \text{ (bobot acak 4)}$$

3. Setelah jarak antara *node* diketahui maka ditentukan nilai minimum dari perhitungan jarak vektor $d_{(j)}$, maka tahap selanjutnya melakukan perubahan bobot .

$$w_{ij}(\text{new}) = w_{ij}(\text{old}) + \alpha [x_i - w_{ij}(\text{old})] \quad \dots 3.5$$

Contoh: dari ke empat vektor jarak tersebut maka dipilih vektor jarak dengan nilai minimum yaitu $d_{(j)}$ pada bobot acak 1.

$$w_{1,1}(\text{new}) = w_{1,1}(\text{old}) + \alpha [x_i - w_{1,1}(\text{old})]$$

$$= 0,3 + 0,05 [-0,228 - 0,3] = 0,2736 \text{ (} w_{1,1} \text{)}$$

Tahap ini dilakukan terus-menerus hingga sampai $w_{1,4}$.

$$w_{1,2} = 0,5 + 0,05 [-0,25 - 0,5] = 0,4625$$

$$w_{1,3} = 0,4 + 0,05 [-0,222 - 0,4] = 0,3689$$

$$w_{1,4} = 0,3 + 0,05 [0,35 - 0,3] = 0,3025$$

4. Pada proses untuk mendapatkan bobot baru memerlukan nilai *learning rate* (α) yaitu $0 \leq \alpha \leq 1$. Nilai *learning rate* pada setiap *epoch* akan berkurang menjadi $\alpha(i+1) = 0,5\alpha$.
 α yang digunakan yaitu 0,05.

Setelah nilai $w_{ij}(\text{new})$ didapatkan maka nilai w_{ij} acak diganti.

Tabel 3.6 Contoh bobot acak baru

Bobot acak	X_1	X_2	X_3	X_4
1	0,2736	0,4625	0,3689	0,3025
2	0,2	0,3	0,6	0,2
3	0,4	0,5	0,3	0,4
4	0,1	0,5	0,8	0,4

5. Kondisi penghentian pengujian dilakukan dengan menghitung selisih antara bobot w_{ij} (*new*) dengan w_{ij} (*old*), apabila nilai w_{ij} hanya berubah sedikit saja, berarti pengujian sudah mencapai konvergensi sehingga dapat dihentikan.

3.8.2 Validasi Cluster

Validasi *cluster* adalah prosedur yang mengevaluasi hasil analisis *cluster* secara kuantitatif dan objektif. Menurut Jain dan Dubes (1988), terdapat tiga pendekatan untuk mengeksplorasi validasi cluster, antara lain:

1. Validasi eksternal, mengevaluasi hasil dari metode *clustering* berdasarkan praspesifikasi struktur yang diterima dari sebuah data yang mencerminkan intuisi pengguna tentang struktur *clustering* dari data.
2. Validasi internal, mengevaluasi hasil *clustering* dalam konsep kuantitatif yang didapat dari data.
3. Validasi relatif, membandingkan sebuah struktur *clustering* dengan struktur *clustering* yang lain yang didapatkan dari metode *clustering* yang sama tetapi nilai parameternya dimodifikasi.

Validasi *cluster* yang peneliti gunakan adalah validasi internal, karena validasi ini didasarkan pada evaluasi hasil *clustering* dalam konsep kuantitatif yang mana data banjir ini juga berupa data kuantitatif. Terdapat beberapa metode atau cara pada validasi internal, diantaranya *connectivity*, nilai *silhouette*, dan indeks *Dunn*.

1. Indeks *Dunn*

Indeks *Dunn* adalah rasio jarak terkecil antara observasi pada *cluster* yang berbeda dengan jarak terbesar pada masing-masing *cluster* data. (Irwansyah dan Faisal, 2015)

Indeks *Dunn* diperoleh dari hasil pembagian antara d_{\min} dengan d_{\max}

$$Dunn = \frac{d_{\min}}{d_{\max}} \quad \dots 3.6$$

dengan,

d_{\min} = jarak terkecil antara observasi pada *cluster* yang berbeda

d_{\max} = jarak terbesar pada masing-masing *cluster* data

2. Indeks *Silhouette*

Indeks *Silhouette* dihitung sebagai derajat kepercayaan dalam proses *clustering* pada suatu pengamatan dengan *cluster* yang dikatakan terbentuk baik bila nilai indeks mendekati 1 dan kondisi sebaliknya jika nilai indeks mendekati angka -1 (Irwansyah dan Faisal, 2015).

$$S_{(i)} = \frac{b_{(i)} - a_{(i)}}{\max(a_{(i)}, b_{(i)})} \quad \dots 3.7$$

dengan,

$a_{(i)}$ = jarak rata-rata antara i dan seluruh pengamatan lainnya

$b_{(i)}$ = jarak rata-rata antara i dengan pengamatan pada *cluster* terdekat

Nilai *sillhouette* dalam rentang -1 hingga 1.

3. Indeks *Connectivity*

$$Conn(C) = \sum_{i=1}^N \sum_{j=1}^L X_{i,nn(i)} \quad \dots 3.8$$

dengan,

$nn_{i(j)}$ = pengamatan tetangga terdekat (*nearest neighbor*) i ke j dan L

$nn_{i(j)}$ = sebagai parameter yang menentukan jumlah tetangga yang berkontribusi pada pengukuran *connectivity*.

3.9 Menginterpretasikan Profil *Cluster*

Menurut Hair dkk (1998), interpretasi *cluster* yaitu langkah dalam analisis *cluster* dengan memeriksa setiap kelompok dari segi variasi *cluster* dalam hal memberikan nama atau menandai dengan suatu label secara tepat yang dapat menggambarkan sifat dari suatu *cluster* atau kelompok. Saat melakukan proses interpretasi, ukuran yang biasa digunakan adalah nilai *centroid cluster*.

Pemprofilan *cluster* meliputi penggambaran karakteristik dari masing-masing *cluster* untuk menjelaskan perbedaan antar *cluster*. Menginterpretasikan dan melakukan profil *cluster* meliputi pengkajian mengenai *centroids* yaitu rata-rata nilai objek yang terdapat dalam *cluster* pada setiap variabel. Nilai *centroid* memungkinkan untuk menguraikan setiap *cluster* dengan cara memberikan suatu nama atau label. (Supranto, 2004).

