BAB II

MATERI PENUNJANG

Untuk menyelesaikan permasalahan clustering diberikan analisa cluster dan graph berarah yangmemberikan dasar yang kuat dalam menyelesaikan BAB III

2.1.Analisa Cluster

Secara umum,jika membicarakan tentang permasalahan clustering harus dianggap bahwa untuk setiap elemen didalam himpunan N diberikan ukuran vektor dengan karakteristik yang digunakan sebagai informasi awal untuk penggelompokkan. Sehingga dibutuhkan asumsi bahwa x_i = ($x_{1i},...x_{zi}$) R^z dengan i $\in$ N. Selanjutnya diberikan definisi jarak antara dua elemen dari N.

Definisi 2.1.

Fungsi bernilai riil d_{ij} dengan i dan j $\in$ N dinamakan jarak euclid yang didefinisikan dalam N, jika memenuhi

(1) $d_{ij} \geq 0$, $d_{ij} = 0$ jika i = j

(2) $d_{ij} = d_{ji}$

(3) $d_{ik} + d_{kj} \geq d_{ij}$

Dengan menganggap beberapa jarak yang biasanya digunakan

1. $d_{ij} = \left[\sum_{k=1}^{z} \left| x_{ik} - x_{jk} \right|^{r} \right]^{1/r}$

2. $d_{ij} = \left[\sum_{k=1}^{z} w_{k} \left| x_{ik} - x_{jk} \right|^{r} \right]^{1/r}$

3. $d_{ij} = (x_i - x_j)` p \, (x_i - x_j)$

dengan p adalah ZxZ matriks simetri semidepinit positip. jarak ini disebut jarak euclid umum, dengan diberikan pasangan karakteristik yang dipelajari. kasus kususnya adalah $d_{ij} = (x_i - x_j)` \sum_{Z}^{-1} (x_i - x_j)$ dengan $\sum_{Z}$ adalah matriks varian-covarian yang menghubungkan karakteristik z yang dipelajari. jarak ini lebih umum dari jarak euclid berat dimana didapat $\sum_{Z} = I_Z$ jarak ini disebut mahalanobis D^2. dengan menotasikan d_{ij} sebagai jarak antara i dan j untuk $i,j \in N$.

jika diberikan N obyek dengan p varibel maka dapat diberikan jajaran matrik $x = (x_{i1}, x_{i2}, \ldots x_{in})$ R^p dimana $i \in N$.matriks x yang diberikan sebagai berikut

$$\mathbf{x} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}_{nxp}$$

Berdasarkan data tersebut dan jarak Euclid didapat Matriks Mahalanobis d_{ij}^2 diberikan dalam bentuk

$$D^2\,(x_j,x_k) = \sum_{i=1}^{n}\left[\frac{x_{ij}-x_j}{sj} - \frac{x_{ik}-x_k}{s_k}\right]^2$$

$$= \frac{1}{s_j^2 - s_k^2}\left[s_j^2 s_k^2 - 2s_j s_k \sum_{i=1}^{n}\left(x_{ij}-x_j\right)\left(x_{ik}-x_k\right) + s_j^2 s_k^2\right]$$

$$= 2\left[1 - \frac{\sum_{i=1}^{n}\left(x_{ij}-x_j\right)\left(x_{ik}-x_k\right)}{s_j s_k}\right]$$

$$= 2(1 - r_{jk})$$

$$d_{ij}^2 = \begin{bmatrix} 0 & d_{12}^2 & \ldots & d_{1n}^2 \\ d_{21}^2 & 0 & \ldots & d_{2n}^2 \\ & & & \\ d_{n1}^2 & d_{n2}^2 & \ldots & 0 \end{bmatrix}$$

Selanjutnya diberikan ilustrasi

Diberikan suatu matriks data yang terdiri atas himpunan subyek(n baris) dan himpunan variabel (p kolom). Sel (i,j) mendaftar gambaran nilai dari subyek (12 jenis mobil) dan variabel (3 sifat:bentuk,ukuran ban,kecepatan). Matriks data disajikan dalam tabel kontingensi

Observasi	X_1	X_2	X_3
1	2	1	2
2	4	3	1
3	6	7	5
4	7	6	3
5	5	4	7
6	8	9	6
7	9	8	7
8	7	10	9
9	10	11	11
10	8	9	8
11	12	11	10
12	9	13	14
ΣX_j	87	92	83
$\overline{X}_j$	7,25	7,67	6,92
S_j	2,620	3,448	3,660

Kemudian dicari matriks $((d_{ij}^2))$ sebagai berikut

$$d_{11}^2 = D^2\left(\hat{X}_{i1}, \hat{X}_{i1}\right) = 2\left(1 - r_{11}\right)$$

$$= 2(1-1)$$

$$= 0$$

$$d_{22}^2 = D^2\left(\hat{X}_{i2}, \hat{X}_{i2}\right) = 2\left(1 - r_{22}\right)$$

$$= 2(1-1)$$

$$= 0$$

$$d_{33}^2 = D^2\left(\hat{X}_{i3}, \hat{X}_{i3}\right) = 2\left(1 - r_{33}\right)$$

$$=2(1-1)$$

$$=0$$

$$d_{12}^2 = D^2\left(\hat{X}_{i1}, \hat{X}_{i2}\right) = 2\left(1 - r_{12}\right)$$

$$=2(1-0{,}886)$$

$$=0{,}228$$

$$d_{13}^2 = D^2\left(\hat{X}_{i1}, \hat{X}_{i3}\right) = 2\left(1 - r_{13}\right)$$

$$=2(1-0{,}749)$$

$$=0{,}502$$

$$d_{23}^2 = D^2\left(\hat{X}_{i2}, \hat{X}_{i3}\right) = 2\left(1 - r_{23}\right)$$

$$=2(1-0{,}889)$$

$$=0{,}222$$

$$d_{ij}^2 = \begin{bmatrix} 0 & d_{12}^2 & d_{13}^2 \\ d_{21}^2 & 0 & d_{23}^2 \\ d_{31}^2 & d_{32}^2 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0{,}228 & 0{,}502 \\ 0{,}228 & 0 & 0{,}222 \\ 0{,}502 & 0{,}222 & 0 \end{bmatrix}$$

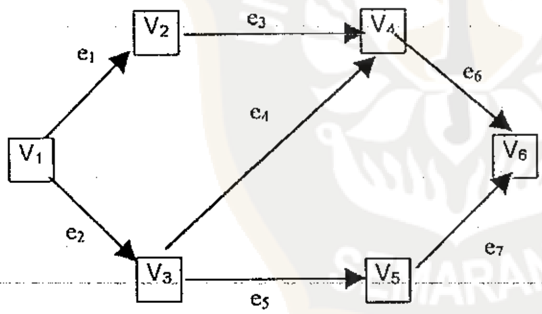
2.2. Graph berarah

Diberikan definisi dari Graph berarah dan aliran dari sebuah jaringan yang digunakan untuk menunjang dalam pembahasan BAB III

Definisi 2.2.

Graph berarah G dinotasikan dengan G(V,E) terdiri dari himpunan titik V dan himpunan garis E yang merupakan pasangan berurutan (x,y) dengan $(x,y) \in E$. selanjutnya x disebut titik awal dan y disebut titik akhir dari garis (x,y). Arah dari suatu garis(x,y) ditunjukkan dengan tanda panah pada garis tersebut.

Contoh:

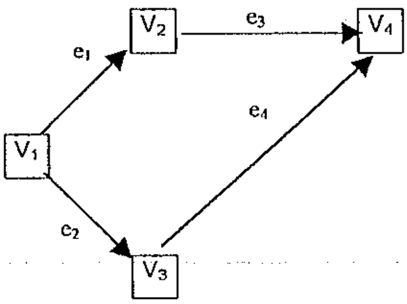


Gambar 1. Suatu graph dengan himpunan titik $\{v_1,v_2,v_3,v_4,v_5,v_6\}$ dan himpunan garis $(e_1,e_2,e_3,e_4,e_5,e_6\}$.

Definisi 2.3.

Graph g dikatakan subgraph dari graph G bila seluruh titik dan garisnya berada dalam G.

contoh



Gambar 2.salah satu subgraph dari graph dalam gambar .1.

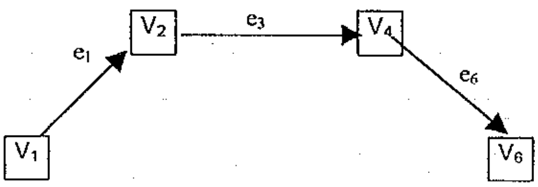
Definisi 2.4.

Path dari graph berarah G(V,E) adalahsuatu subgraph dari graph berarah G(V,E) yang terdiri dari barisan titik dan garis {v_1-e_1-v_2-e_2-…-v_{r-1}-e_{r-1}-v_r} tanpa pengulangan garis.

Untuk setiap $1 \leq k \leq r-1$, $e_k = (v_k, v_{k+1}) \in E$ disebut garis maju,dan $e_k = (v_{k+1}, v_k) \in E$ disebut garis balik.

Jika v_1=v_r maka disebut path tertutup atau cycle dan jika $v_1 \neq v_r$ maka disebut path terbuka.

Contoh:

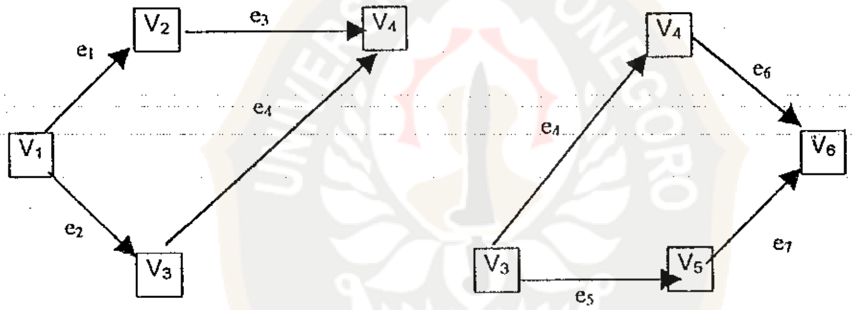


Gambar 3. Path terbuka dari graph dalam gambar1{v_1-e_1-v_2-e_3-v_4-e_6-v_6}.

Definisi 2.5.

Graph G_1 (V_1,E_1)dan G_2 (V_2,E_2) adalah subgraph dari G_3 ,gabungan dari G_1 dan G_2 ditulis sebagai $G_3 = G_1 \cup G_2$ dengan himpunan titiknya $V_3 = V_1 \cup V_2$ dan himpunan garisnya $E_3 = E_1 \cup E_2$. Selanjutnya $G_1 \cup G_2$ disebut jumlah dari gaph G_1 dan G_2 yang mempunyai sifat $G_1 \cup G_2 = G_2 \cup G_1$.

Contoh

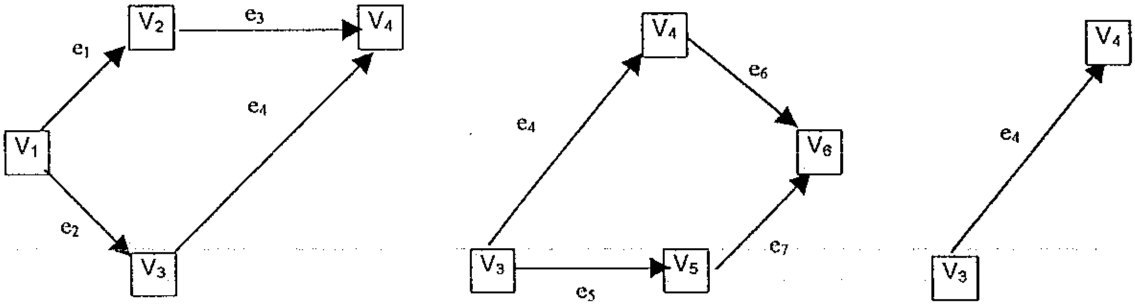


Gambar4. subgraph dari jaringan dalam gambar1gabungan dari 2 subgraph membentuk graph dalam gambar 1.

Definisi 2.6.

Graph G_4 adalah subgraph dari G_1 (V_1,E_1) dan G_2 (V_2,E_2) sehingga irisan dari G_1 dan G_2 yang ditulis dengan $G_4 = G_1 \cap G_2$ dengan himpunan titiknya $V_4 = V_1 \cap V_2$ dan himpunan garisnya $E_4 = E_1 \cap E_2$. Selanjutnya $G_1 \cap G_2$ disebut irisan dari G_1 dan G_2 yang mempunyai sifat $G_1 \cap G_2 = G_2 \cap G_1$.

Contoh: G_1 G_2 G_4



Gambar 5. Irisan dari dua buah graph.

Definisi 2.7.

Suatu pola aliran {f(x,y)} dikatakan fisibel dalam G(V,E,c,f) dan mempunyai nilai aliran f_s dari titik sumber s ke titik terminal t jika untuk setiap $x \in V$ terpenuhi

$$\sum_y f(x,y) - \sum_y f(y,x) = f_{st}, x = s$$

$$= 0, x \neq s, t$$

$$= -f_{st}, x = t$$

$$c(x,y) \geq f(x,y) \geq 0, (x,y) \in E$$