

**ANALISIS KELOMPOK (CLUSTER) DENGAN
MENGUNAKAN METODE SELF ORGANIZING MAPS
PADA DATA WINES**

SKRIPSI

**oleh:
SUSAN ADI PUTRA
0410950053-95**



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2009**

**ANALISIS KELOMPOK (CLUSTER) DENGAN
MENGUNAKAN METODE SELF ORGANIZING MAPS
PADA DATA WINES**

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Sains dalam bidang Statistika

oleh:
SUSAN ADI PUTRA
0410950053-95



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2009**

LEMBAR PENGESAHAN SKRIPSI

**ANALISIS KELOMPOK (*CLUSTER*) DENGAN
MENGUNAKAN METODE *SELF ORGANIZING MAPS*
PADA DATA WINES**

oleh:
SUSAN ADI PUTRA
0410950053-95

Setelah dipertahankan di depan Majelis Penguji
pada tanggal 20 April 2009
dan dinyatakan memenuhi syarat untuk memperoleh gelar
Sarjana Sains dalam bidang Statistika

Pembimbing I

Pembimbing II

Prof. Dr. Ir. Waego H. Nugroho
NIP. 130 074 146

Adji Achmad R. F. SSi., MSc.
NIP. 132 233 148

Mengetahui,
Ketua Jurusan Matematika
Fakultas MIPA Universitas Brawijaya

Dr. Agus Suryanto, MSc.
NIP. 132 126 049

LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama : Susan Adi Putra
NIM : 0410950053-95
Jurusan : Matematika
Penulisan skripsi berjudul : **ANALISIS KELOMPOK
(CLUSTER) DENGAN
MENGUNAKAN METODE
SELF ORGANIZING MAPS PADA
DATA WINES**

Dengan ini menyatakan bahwa :

1. Isi dari skripsi yang saya buat adalah benar-benar karya sendiri dan tidak menjiplak karya orang lain, selain nama-nama yang termaktub di isi dan tertulis di daftar pustaka dalam skripsi ini.
2. Apabila di kemudian hari ternyata skripsi yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung segala risiko.

Demikian pernyataan ini saya buat dengan segala kesadaran.

Malang, 20 April 2009
Yang menyatakan,

Susan Adi Putra
NIM. 0410950053-95

ANALISIS KELOMPOK (*CLUSTER*) DENGAN MENGUNAKAN METODE *SELF ORGANIZING MAPS* PADA DATA *WINES*

ABSTRAK

Analisis kelompok (*cluster*) merupakan suatu metode dalam analisis peubah ganda yang digunakan untuk mengelompokkan sejumlah objek pengamatan yang relatif homogen menjadi kelompok yang lebih kecil berdasarkan karakteristik peubah. Ada dua teknik pengelompokan, yaitu teknik pengelompokan hirarki dan teknik pengelompokan nonhirarki. Teknik pengelompokan nonhirarki yang sering digunakan adalah metode *K-means*. Selain metode *K-means* terdapat teknik pengelompokan nonhirarki lain yang dapat digunakan yaitu metode *Self Organizing Maps* (SOM). SOM adalah teknik pengelompokan yang digunakan untuk memvisualisasikan atau menggambarkan objek pengamatan dengan mereduksi dimensi objek pengamatan kelebihan ini tidak terdapat pada *K-means*. Tujuan dari SOM adalah untuk memvisualisasikan objek pengamatan dengan mengelompokkan objek pengamatan tersebut ke dalam *map-map* yang berbentuk *hexagonal* berdasarkan warna yang menyusun *map-map* dan untuk menentukan tingkat *learning* serta *layer maps* yang optimal. Data yang digunakan merupakan data hasil penelitian pada *wines*. Setelah menghitung kedekatan jarak dengan menggunakan jarak Euclid, objek pengamatan yang mempunyai jarak yang saling berdekatan digabungkan dalam satu kelompok. Berdasarkan nilai *learning* dan *layer hexagonal* yang digunakan dalam penelitian ini dapat disimpulkan bahwa nilai *learning* yang paling kecil dapat memvisualisasikan kelompok yang terbentuk secara jelas dengan tingkat warna yang berbeda dan *layer maps* yang optimal digunakan dalam pengelompokan objek pengamatan adalah *layer maps hexagonal 5x5*.

Kata kunci: *Cluster, Self Organizing Maps, Wines*

CLUSTERING ANALYSIS USING SELF ORGANIZING MAPS METHOD IN WINES DATA

ABSTRACT

Cluster analysis is the method of analysis of multivariate used to cluster some object observations which is relatively similar in characteristics to be smaller clusters based on each variable of characteristics. There are two clustering analysis to cluster objects in clustering analysis which are hierarchy clustering techniques and non hierarchy clustering techniques. The non hierarchy clustering techniques which is often used is K-means method. Besides K-means methods there is another technique in non hierarchy clustering techniques which can be used. It is called Self Organizing Maps (SOM). SOM is a clustering techniques used to visualize or draw object observations with the reduction of the dimension of the object observations. Goal from SOM is visualizes object observation with clustering object observation in maps based on the different of color that is able to create the maps and to determine optimal learning rate with layer maps. Data that used in this research is the result of wines observation. After account the similarity distance using Euclidean distance, object observation that have closest distance are join in one cluster. According to learning rate and hexagonal layer that used in this research it can be can conclude that the smallest learning rate can visualizes number of cluster taht created clearly with different colour intensity and optimal layer maps hexagonal that used in clustering object observation is 5x5 layer maps hexagonal.

Keywords: Cluster, Self Organizing Maps, Wines

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul Analisis Kelompok dengan Menggunakan Metode *Self Organizing Maps* Pada Data *Wines* sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains dalam bidang Statistika. Oleh karena itu, penulis mengucapkan rasa hormat dan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Waego Hadi Nugroho, selaku Dosen Pembimbing I atas bimbingan yang telah diberikan kepada penulis selama penyusunan skripsi ini.
2. Bapak Adji Achmad Rinaldo F., SSi., MSc., selaku Dosen Pembimbing II atas kesabaran dan arahan yang diberikan kepada penulis selama penyusunan skripsi ini.
3. Ibu Ir. Soepraptini, MSc., Ibu Suci Astutik, SSi., MSi. dan Ibu Dra. Ani Budi Astuti, MSi. selaku Dosen Penguji.
4. Bapak dan Ibu Dosen Statistika atas didikan selama kuliah hingga penulis bisa menyelesaikan kuliah.
5. Papa, Mama, Adek Kiki, Adek Dewi, Adek Putri, Ika Afriniawati, Mas Demi dan keluarga besar yang senantiasa mendoakan dan membantu penulis mencapai yang terbaik.
6. Teman-teman Program Studi Statistika 2002, 2003, 2004 dan 2005 yang telah memberikan dukungan dan bantuan.
7. Semua pihak yang tidak dapat penulis sebutkan satu per satu yang telah banyak membantu dan memberikan dorongan selama penulisan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan mengingat keterbatasan kemampuan penulis. Untuk itu, dengan segala kerendahan hati penulis mengharap kritik dan saran. Akhirnya penulis berharap semoga skripsi ini dapat bermanfaat bagi pembaca.

Malang, 20 April 2009

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	viii
DAFTAR TABEL	x
DAFTAR LAMPIRAN	xi
BAB I PENDAHULUAN	
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah	2
1.4. Tujuan Penelitian.....	2
1.5. Manfaat Penelitian.....	2
BAB II TINJAUAN PUSTAKA	
2.1. Analisis Kelompok (<i>Cluster</i>).....	3
2.1.1. Analisis Kelompok Hierarki	3
2.1.2. Analisis Kelompok Non Hierarki	4
2.2. Jaringan Syaraf Tiruan.....	6
2.3. <i>Self Organizing Maps</i> (SOM).....	9
2.3.1. Pembelajaran (<i>Supervised</i>) Learning.....	9
2.3.2. Jaringan SOM Kohonen	10
2.3.3. Visualisasi <i>Self Organizing Maps</i> (SOM)	13
2.3.4. Warna Dalam <i>Self Organizing Maps</i> (SOM).....	14
BAB III METODOLOGI	
3.1. Data.....	17
3.2. Metode.....	19
BAB IV HASIL DAN PEMBAHASAN	
4.1. Analisis Kelompok Non Hierarki	21
4.2. Visualisasi Data <i>Wines</i> dengan SOM.....	22
BAB V KESIMPULAN DAN SARAN	
5.1. Kesimpulan.....	39
5.2. Saran.....	39
DAFTAR PUSTAKA	41
LAMPIRAN	43

DAFTAR GAMBAR

	Halaman
Gambar 2.1. Struktur Dasar Jaringan Syaraf Tiruan dan Struktur Sederhana Sebuah <i>Neuron</i>	8
Gambar 2.2. Sistem <i>Learning Supervised Self Organizing Maps</i> (SOM)	9
Gambar 2.3. Kohonen <i>Network</i>	10
Gambar 2.4. Contoh dari Representasi Gambar Skala Jarak 5x5 <i>hexagonal</i>	14
Gambar 2.5. Skala Warna RGB (<i>Red, Green, dan Blue</i>)	15
Gambar 3.1. Diagram Alir Analisis Kelompok NonHierarki <i>Self Organizing Maps</i> (SOM).....	20
Gambar 4.1. <i>Mapping Plot</i> Data <i>Wines</i> 3x3 <i>Layer Hexagonal</i>	22
Gambar 4.2. Visualisasi SOM Distribusi Objek Pengamatan dalam Kelompok 3x3 <i>Layer Hexagonal</i>	22
Gambar 4.3. Visualisasi SOM Konsentrasi Bahan Penyusun <i>Wines</i> 3x3 <i>Layer Hexagonal</i>	23
Gambar 4.4. Visualisasi SOM Kelompok yang Terbentuk 3x3 <i>Layer Hexagonal</i>	23
Gambar 4.5. <i>Mapping Plot</i> Data <i>Wines</i> 4x4 <i>Layer Hexagonal</i>	24
Gambar 4.6. Visualisasi SOM Distribusi Objek Pengamatan dalam Kelompok 4x4 <i>Layer Hexagonal</i>	25
Gambar 4.7. Visualisasi SOM Konsentrasi Bahan Penyusun <i>Wines</i> 4x4 <i>Layer Hexagonal</i>	25
Gambar 4.8. Visualisasi SOM Kelompok yang Terbentuk 4x4 <i>Layer Hexagonal</i>	26
Gambar 4.9. <i>Mapping Plot</i> Data <i>Wines</i> 5x5 <i>Layer Hexagonal</i>	27
Gambar 4.10. Visualisasi SOM Distribusi Objek Pengamatan dalam Kelompok 5x5 <i>Layer Hexagonal</i>	27

Gambar 4.11. Visualisasi SOM Konsentrasi Bahan Penyusun <i>Wines</i> 5x5 <i>Layer Hexagonal</i>	28
Gambar 4.12. Visualisasi SOM Kelompok yang Terbentuk 5x5 <i>Layer Hexagonal</i>	28
Gambar 4.13. <i>Mapping Plot</i> Data <i>Wines</i> 6x6 <i>Layer Hexagonal</i>	29
Gambar 4.14. Visualisasi SOM Distribusi Objek Pengamatan dalam Kelompok 6x6 <i>Layer Hexagonal</i>	30
Gambar 4.15. Visualisasi SOM Konsentrasi Bahan Penyusun <i>Wines</i> 6x6 <i>Layer Hexagonal</i>	30
Gambar 4.16. Visualisasi SOM Kelompok yang Terbentuk 6x6 <i>Layer Hexagonal</i>	31
Gambar 4.17. Visualisasi SOM dengan Tingkat <i>Learning</i> 0.1 5x5 <i>Layer Hexagonal</i>	32
Gambar 4.18. Visualisasi SOM dengan Tingkat <i>Learning</i> 0.9 5x5 <i>Layer Hexagonal</i>	33
Gambar 4.19. Visualisasi SOM dengan Tingkat <i>Learning</i> 0.7 5x5 <i>Layer Hexagonal</i>	33
Gambar 4.20. Visualisasi SOM dengan Tingkat <i>Learning</i> 0.5 5x5 <i>Layer Hexagonal</i>	34
Gambar 4.21. Visualisasi SOM dengan Tingkat <i>Learning</i> 0.3 5x5 <i>Layer Hexagonal</i>	34
Gambar 4.22. Visualisasi SOM dengan Tingkat <i>Learning</i> 0.1 5x5 <i>Layer Hexagonal</i>	35

DAFTAR TABEL

	Halaman
Tabel 2.1. Rumus Pergeseran Warna RGB.....	15
Tabel 4.1. Hasil Pengelompokan Objek Pengamatan dengan SOM.....	35
Tabel 4.2. Nilai Tengah Peubah <i>Wines</i> Hasil <i>Training</i>	36

UNIVERSITAS BRAWIJAYA



DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data <i>Wines</i>	43
Lampiran 2. Matriks Jarak Antar <i>Neuron 5x5 Hexagonal</i>	44
Lampiran 3. Indeks Pengelompokan Objek Pengamatan Dengan SOM.....	45
Lampiran 4. Indeks Kelompok Yang Terbentuk.....	46
Lampiran 5. Hasil <i>Training</i> dengan SOM	47
Lampiran 6. Matriks X_{ik}	49
Lampiran 7. Macro R untuk Data <i>Wines</i>	50



BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis kelompok (*cluster*) merupakan suatu metode dalam analisis peubah ganda yang digunakan untuk mengelompokkan sejumlah objek pengamatan yang relatif homogen menjadi kelompok yang lebih kecil berdasarkan karakteristik peubah (Johnson & Wichern, 1992).

Prinsip utama dari analisis kelompok adalah mengelompokkan sejumlah objek pengamatan yang homogen dalam suatu kelompok yang lebih sederhana. Kehomogenan ini didasarkan pada ukuran kesamaan antar objek pengamatan. Ukuran kedekatan antar objek pengamatan yang digunakan tergantung dari jenis data yang dianalisis. Jika data yang dianalisis bersifat kuantitatif (metrik), ukuran kesamaan yang digunakan adalah ukuran jarak. Jika data yang dianalisis bersifat kualitatif (nonmetrik), maka ukuran kesamaan yang digunakan adalah ukuran asosiasi (Hartigan, 1975).

Ada dua teknik pengelompokan yang dapat digunakan dalam mengelompokkan objek dalam analisis kelompok, yaitu teknik pengelompokan hirarki dan teknik pengelompokan nonhirarki. Analisis kelompok hirarki dimulai dengan mengelompokkan dua atau lebih objek yang mempunyai kesamaan paling dekat. Sedangkan analisis kelompok nonhirarki dimulai dengan menentukan terlebih dahulu jumlah kelompok yang diinginkan. Perbedaan antara kedua metode pengelompokan ini terletak pada penentuan jumlah kelompok yang dihasilkan. Pada teknik pengelompokan hirarki banyaknya kelompok yang dihasilkan belum diketahui. Sedangkan pada teknik pengelompokan nonhirarki banyaknya kelompok ditentukan terlebih dahulu (Johnson & Wichern, 1992).

Pada teknik pengelompokan nonhirarki, metode yang sering digunakan yaitu metode *K-means*. Metode *K-means* adalah suatu metode yang mengelompokkan sejumlah objek ke dalam kelompok objek yang lebih sederhana berdasarkan nilai rata-rata objek. Selain metode *K-means* dalam analisis kelompok nonhirarki terdapat metode lain yang dapat digunakan untuk pengelompokan, yaitu metode *Self Organizing Maps* (SOM). SOM adalah teknik pengelompokan dengan memvisualisasikan atau menggambarkan

objek pengamatan dengan mereduksi objek pengamatan. SOM adalah salah satu teknik jaringan syaraf tiruan yang ditemukan untuk menganalisis sekelompok data yang kompleks. Penggunaan SOM tidak hanya diterapkan dalam bidang teknik tetapi diterapkan di area lain yaitu kedokteran, pertanian dan ilmu sosial termasuk untuk menggambarkan demografi suatu daerah (Tokukata, 1999).

Kelebihan metode SOM dibandingkan dengan metode *K-means* adalah di dalam SOM hasil pengelompokan objek pengamatan dapat divisualisasikan dalam bentuk gambar dengan warna tertentu dan visualisasi objek pengamatan ini tidak terdapat pada metode *K-means*. Visualisasi atau penggambaran dari metode SOM inilah yang akan menjadi bahasan dalam skripsi ini.

1.2 Rumusan Masalah

Rumusan masalah skripsi ini adalah bagaimana menerapkan metode *Self Organizing Maps* (SOM) dalam memvisualisasikan atau menggambarkan pengelompokan data?

1.3 Batasan Masalah

Batasan masalah dalam Skripsi ini adalah :

1. Banyaknya *layer maps* 3x3; 4x4; 5x5 dan 6x6 (*hexagonal*/segi enam).
2. Parameter $\alpha_0 = 0.1; 0.3; 0.5; 0.7; 0.9$ dan iterasi = 100.
3. Dimensi data adalah 2.
4. Banyaknya kelompok adalah 3 kelompok.

1.4 Tujuan

Tujuan dari Skripsi ini adalah:

1. Memvisualisasikan hasil pengelompokan data *wines*.
2. Menentukan nilai *learning* dan *layer maps* optimal dalam pengelompokan objek pengamatan data *wines*..

1.5 Manfaat

Manfaat penelitian ini adalah metode *Self Organizing Maps* (SOM) sebagai alternatif metode pengelompokan (*cluster*) yang dapat memvisualisasikan hasil pengelompokan objek pengamatan pada data *wines* dalam bentuk *layer hexagonal* dengan perbedaan warna pada masing-masing kelompok yang terbentuk.

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Kelompok (*Cluster*)

Analisis kelompok merupakan suatu teknik yang digunakan untuk mengelompokkan objek pengamatan menjadi beberapa kelompok. Objek pengamatan dalam setiap kelompok cenderung homogen satu sama lain dan tidak homogen terhadap kelompok lainnya. Tujuan pokok dari analisis kelompok adalah untuk mengelompokkan objek pengamatan menjadi beberapa kelompok yang lebih sederhana berdasarkan tingkat kehomogenan objek pengamatan. Analisis kelompok sering digunakan karena tidak ada hipotesis parameter yang diuji. Jadi pengelompokan hanya berdasarkan fakta dari data tersebut (Matjijk *et al*, 2002).

Vesanto (2002) menyatakan bahwa proses dari analisis kelompok (*cluster*) adalah pengelompokan data yang dilakukan dengan dua metode, yaitu metode hirarki dan metode nonhirarki. Pada metode nonhirarki telah ditentukan jumlah kelompok terlebih dahulu. Sedangkan metode hirarki digunakan apabila jumlah kelompok ditentukan berdasarkan hasil analisis. Metode pengelompokan nonhirarki digunakan apabila pada awal analisis kurang adanya informasi mengenai karakteristik objek pengamatan yang akan dikelompokkan.

2.1.1 Analisis Kelompok Hirarki

Pada analisis pengelompokan hirarki dianggap bahwa mula-mula tiap-tiap objek merupakan suatu kelompok tersendiri, lalu dua objek atau kelompok yang terdekat digabungkan sehingga menjadi suatu kelompok yang lebih kecil (Johnson & Wichern, 1992).

Teknik analisis hirarki terdiri atas dua metode, yaitu (1) *agglomerative* (penggabungan) di mana masing-masing objek dianggap satu kelompok kemudian antar kelompok yang jaraknya berdekatan bergabung menjadi satu kelompok dan (2) *divisive* (pemecahan) yaitu pada awalnya semua objek berada dalam satu kelompok kemudian sifat paling beda dipisahkan dan membentuk satu kelompok yang lain (Johnson & Wichern, 1992).

Dalam proses penggabungan kelompok dengan metode hirarki selalu disertai dengan perbaikan matriks jarak atau matriks kesamaan. Metode perbaikan jarak adalah sebagai berikut:

1. Metode pautan lengkap (*Complete Linkage*), yaitu memaksimumkan jarak antar pasangan kelompok yang digabungkan. Metode ini mengelompokkan dua objek yang memiliki jarak terjauh.

Duran & Odell (1974) menyebut metode ini dengan *further neighbour distance* karena kelompok-kelompok digabungkan berdasarkan jarak terjauh antara kelompok-kelompok yang ada.

2. Metode pautan rata-rata (*Average Linkage*), yaitu meminimumkan rata-rata jarak antar pasangan kelompok yang digabungkan. Metode ini mengelompokkan dua objek yang mempunyai jarak terdekat terlebih dahulu.

3. Metode pautan tunggal (*Single Linkage*), yaitu meminimumkan jarak antar kelompok yang digabungkan. Kelompok ini dua objek yang terdekat terlebih dahulu.

Menurut William & Lambert (1966) menyebut metode ini dengan *nearest neighbour distance* karena pada tiap tahap, kelompok-kelompok digabungkan berdasarkan jarak terpendek antara kelompok yang terbentuk.

Metode Ward (*Ward's Method*) yaitu meminimumkan varians total sehingga diperoleh kelompok yang homogen. Pada metode ini, jarak antara dua kelompok yang terbentuk adalah *sum of square* diantara dua kelompok tersebut.

4. Metode *Centroid* (*Centroid Method*). Pada metode ini, jarak antara dua kelompok adalah jarak diantara *centroid* pada kelompok itu. *Centroid* adalah rata-rata jarak yang ada pada sebuah kelompok yang didapat dengan melakukan rata-rata pada semua anggota kelompok.

Duran & Odell (1974)

2.1.2 Analisis Kelompok NonHirarki

Pada analisis pengelompokan nonhirarki ditentukan dahulu jumlah kelompok yang diinginkan sehingga sifat pengelompokannya

tidaklah alamiah karena dikondisikan untuk jumlah kelompok tertentu. Proses dimulai dengan nilai k-kelompok yang merupakan pusat kelompok dengan cara acak dari data. Setelah itu proses kelompok dilakukan tanpa melalui proses hirarki (Johnson & Wichern, 1992).

Dalam analisis nonhirarki terdapat teknik dalam pembentukan kelompok. Berikut adalah teknik pembentukan kelompok dalam analisis nonhirarki:

1. *Sequential Threshold*
Metode ini melakukan pengelompokan diawali dengan memilih satu objek sebagai dasar nilai awal kelompok, kemudian semua objek yang mempunyai jarak terdekat dengan kelompok ini digabung lalu dipilih kelompok kedua dan semua objek yang mempunyai kemiripan/homogen dimasukkan dalam kelompok ini. Demikian seterusnya hingga terbentuk keseluruhan objek yang ada di dalamnya.
2. *Parallel Threshold*
Metode ini hampir sama dengan metode *Sequential Threshold* hanya saja pemilihan terhadap beberapa objek awal sekaligus dan kemudian menggabungkan kelompok secara bersamaan.
3. *Optimizing Partitioning*
Merupakan pengembangan dari kedua metode di atas dengan melakukan optimasi pada objek yang ditukar untuk kelompok lainnya dengan pertimbangan kriteria optimasi.

(Sharma, 1996)

Metode K-Means dan SOM merupakan bagian dari teknik *Optimizing Partitioning*. Metode K-Means adalah suatu metode yang mengelompokkan sejumlah objek ke dalam kelompok objek yang lebih sederhana berdasarkan nilai rata-rata objek.

Mattjik *et. al.*(2002) menjelaskan bahwa proses pengelompokan K-Means terdiri atas 3 tahapan, yaitu:

1. Bagi individu-individu ke dalam k-kelompok awal, kemudian hitung rata-rata awal.
2. Masukkan tiap-tiap individu ke suatu kelompok berdasarkan rata-rata terdekat. Hitung kembali rata-rata kelompok baru yang terbentuk.
3. Ulangi langkah 2) sampai tidak ada lagi perpindahan individu ke kelompok lain.

Algoritma pengelompokan dengan menggunakan metode *K-Means* digunakan untuk meminimalkan jumlah kuadrat jarak di antara *input* vektor. Sedangkan Metode *Self Organizing Maps* (SOM) adalah teknik pengelompokan dengan memvisualisasikan atau menggambarkan objek pengamatan dengan mereduksi objek pengamatan. Metode *Self Organizing Maps* (SOM) merupakan bagian dari jaringan syaraf tiruan yang terdiri atas *input* dan *output layer* (Kohonen, 1989).

2.2 Jaringan Syaraf Tiruan

Geugan (1999) menjelaskan bahwa *neuron* (sel syaraf tiruan) mengirim *signal* elektro kimia melalui jalur *neural*. Setiap *neuron* menerima *signal input* dari *neuron* melalui *synapse* (perantara neuron). Beberapa *input* membangkitkan *neuron* dan yang lainnya menghambat, ketika efek kumulatif melebihi batas maka *neuron* akan menembak dan mengirim *signal* ke *neuron* lainnya. Model jaringan syaraf tiruan mempunyai karakteristik sama dengan *neuron* biologi seperti Gambar 2.1. Setiap *neuron* menerima satu set *input*. Setiap *input* akan dikurangkan dengan bobot masing-masing kemudian akar dari total pengurangan tersebut dikuadratkan. Hal tersebut dapat dinotasikan sebagai berikut:

$$d_i = \|X_i - W_j\| = \sqrt{\sum_{k=1}^p (X_{ik} - W_{jk})^2} \quad 2.1$$

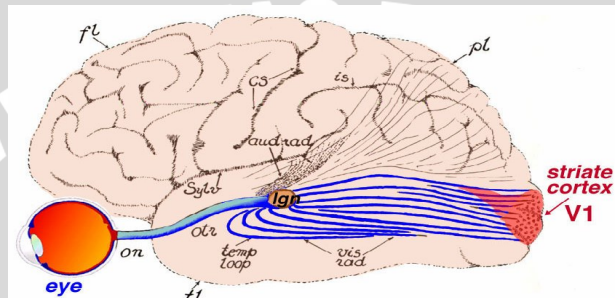
di mana:

- d_i : jarak objek pengamatan ke- i
- X_i : input objek pengamatan yang berukuran $i \times k$
- W_j : pembobot objek pengamatan yang berukuran $j \times k$
- i : 1,2,...,n
- k : 1,2,...,p
- j : 1,2,3

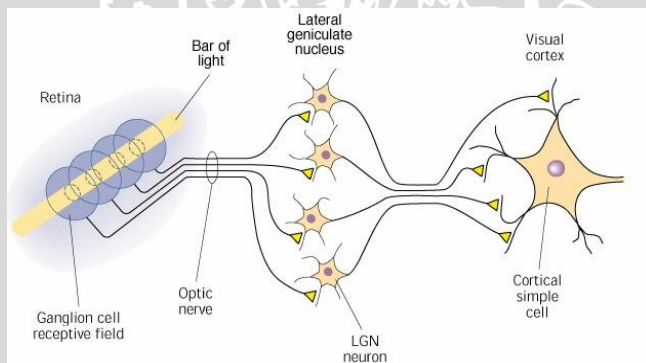
Nilai pembobot W_{jk} merupakan nilai acak antara 0 sampai 1 yang didapatkan dari tabel bilangan acak berdasarkan baris dan kolom yang ada pada table bilangan acak tersebut.

Signal input kemudian diproses dengan menggunakan fungsi aktivasi untuk menghasilkan *signal output*. Proses di atas dapat digambarkan dengan urutan seperti berikut:

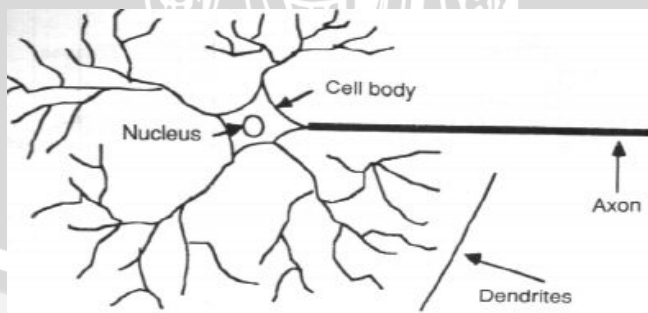
a)



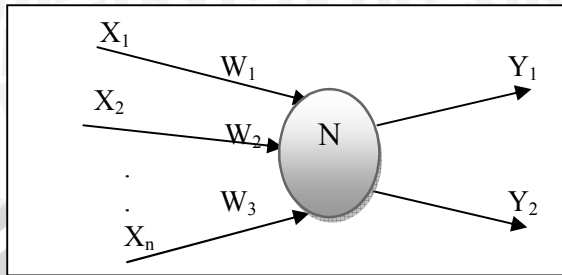
b)



c)



d)



(Kohonen, 1995)

Gambar 2.1 Struktur Dasar Jaringan Syaraf Tiruan dan Struktur Sederhana Sebuah *Neuron*.

Pada Gambar 2.1.a) dapat dijelaskan bahwa *input* yang berasal dari luar (rangsangan cahaya) masuk dalam mata melalui jaringan kemudian diproses oleh selaput otak untuk memproses rangsangan kemudian diteruskan menuju otak kecil untuk menggerakkan syaraf agar menghasilkan informasi yang jelas berupa gambar yang jelas mengenai rangsangan cahaya (*output*).

Gambar 2.1.b) menjelaskan bagaimana rangsangan diproses dalam suatu jaringan syaraf. Rangsangan yang diterima oleh retina diteruskan melalui optik yang kemudian diproses melalui *neuron* setelah rangsangan diproses kemudian divisualisasikan melalui suatu *neuron* yaitu *visual cortex* sehingga rangsangan dapat digambarkan secara jelas setelah melalui proses ini.

Gambar 2.1.c) merupakan gambar dari satu jaringan syaraf yang terdiri dari *axon*, *dendrites*, dan *soma*. *Axon* berfungsi untuk meneruskan rangsangan, *dendrites* menerima *signal* dari *neuron* lain. *Signal* tersebut berupa *impuls* elektrik kimiawi yang dimodifikasi (diperkuat/diperlemah) di celah *synapse*. *Soma* berfungsi menjumlahkan semua *signal-signal* yang masuk.

Gambar 2.1.d) adalah proses elemen dengan dua *output* di mana banyaknya peubah adalah X_i dengan bobot W_j yang dianalisis menggunakan satu *neuron* (N). Sebelum dilakukan analisis pada jaringan syaraf tiruan, maka dilakukan *learning*. *Learning* dalam jaringan syaraf tiruan adalah bagaimana menemukan himpunan bobot dari *input*. Dengan memberikan pasangan himpunan *input* dan himpunan *output* yang diinginkan ke dalam jaringan syaraf tiruan,

maka jaringan syaraf tiruan memodifikasi nilai bobot dari *input* dan *output* yang dipasangkan tersebut.

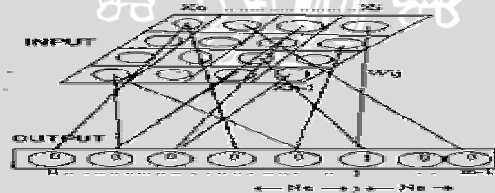
Proses *learning* pada jaringan syaraf tiruan bisa berupa *supervised* atau *unsupervised*. Jaringan syaraf tiruan dikatakan *learning supervised* jika target *output* (hasil pengelompokan) yang diinginkan telah diketahui seperti pada data *wines* di mana banyaknya kelompok ditentukan terlebih dahulu sebanyak 3 kelompok. Sedangkan *learning unsupervised* target *output* belum diketahui, target *output* ditentukan pada waktu proses *learning*. Sasaran *unsupervised* adalah mengelompokkan unit yang mempunyai kesamaan dalam jarak tertentu (Guegan, 1999).

2.3 Self Organizing Maps (SOM)

SOM adalah salah satu jenis jaringan syaraf tiruan yang digunakan untuk memvisualisasikan atau menggambarkan data yang berasal dari dimensi tinggi ke dimensi dua atau satu dengan tetap mempertahankan komposisi data tersebut. Seperti dalam jaringan syaraf tiruan, SOM mempunyai dua model yaitu *training* dan *mapping*. Proses *training* adalah membentuk *map* (gambar) menggunakan *input* data. Komponen SOM berupa *neuron*. SOM menggunakan proses *learning* untuk membuat perbedaan warna dalam jaringan sebagai respon pengelompokan *input* (Ultsch, 2007).

2.3.1 Pembelajaran (Learning) Supervised

Sistem *learning supervised* adalah sistem pembelajaran di mana pada jaringan syaraf tiruan diberikan set *input* dengan memberikan set target yang akan dihasilkan. Sistem ini akan membedakan *input* ke beberapa kelompok (*cluster*) berdasarkan tingkat kesamaan warna yang dimiliki oleh himpunan *input* tersebut (Dunn, 1982). Berikut adalah gambar mengenai sistem *learning supervised*.



Gambar 2.2 Sistem *Learning Supervised Self Organizing Maps* (SOM)

Gambar 2.2 terdapat *input* X_i dalam bentuk *input layer* yang berbentuk *rectangular* (persegi) 4×4 yang diberi pembobot W_j untuk di *training* dengan target *output* yang ditentukan sebanyak 8 buah. Proses *training* dilakukan menggunakan nilai *learning* tertentu dengan iterasi tertentu sampai nilai X_i dengan pembobot W_j menghasilkan jarak yang minimum. Tingkat *learning* (pembelajaran) dinotasikan sebagai $\alpha(t)$ yang nilainya berubah seiring dengan bertambahnya nilai iterasi. Nilai *learning* digunakan untuk mengetahui tingkat ketajaman dari warna dalam visualisasi hasil pengelompokkan objek pengamatan. Persamaan dari tingkat *learning* adalah sebagai berikut:

$$\alpha(t) = \alpha_0 \left(1 - \frac{t}{10^4} \right) \quad 2.2$$

di mana:

$\alpha(t)$: Tingkat *learning* dengan iterasi ke- t

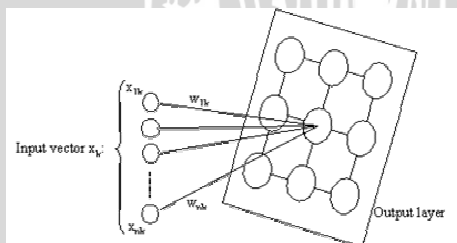
α_0 : 0.1

t : 1,.....,100

2.3.2 Jaringan SOM Kohonen

Jaringan SOM (*Self Organizing Map*) Kohonen merupakan salah satu model jaringan syaraf yang menggunakan metode *learning supervised* (Tokukata, 1999).

Jaringan SOM Kohonen terdiri dari dua lapisan (*layer*), yaitu lapisan *input* dan lapisan *output*. Setiap *neuron* dalam lapisan *input* terhubung dengan setiap *neuron* pada lapisan *output*. Setiap *neuron* dalam lapisan *output* merepresentasikan kelas dari *input* yang diberikan.



Gambar 2.3 Kohonen Network

Kohonen (2001) menjelaskan bahwa gambaran dasar dari SOM adalah fokus pada pembelajaran (*learning*) analisis eksploratori *unsupervised*. Bagaimanapun juga SOM dapat juga digunakan dalam pembelajaran *supervised*. Sebagai informasi tambahan adalah terdapat informasi kelas yang bisa dimodelkan sebagai variabel dependen agar prediksi dapat diperoleh. Data asli biasanya diindikasikan dengan X dan tambahan informasi dengan Y. Untuk membuat *training* SOM adalah menggunakan rangkaian matriks X dan Y.

Data yang berasal dari p-deret percobaan dapat direpresentasikan ke dalam nxp matriks data *wines*, di mana setiap baris mengandung p elemen vektor *wines*(Aas, 2001).

Ultsch (1993) menggunakan pendekatan yang fleksibel dimana jarak dalam ruang X dan Y dapat dihitung secara terpisah. Keduanya diskalakan sehingga persamaan jarak maksimal secara keseluruhan jumlah dari kedua bobotnya. Berikut adalah rumus untuk jarak dalam ruang X dan Y menggunakan jarak Euclid dengan tingkat *learning* tertentu.

$$D(X_{ik}, Y_{jk}) = \alpha_0 D(X_{ik}) + (1 - \alpha_0) D(Y_{jk}) \quad 2.3$$

di mana:

- $D(X_{ik}, Y_{ji})$: kombinasi jarak objek pengamatan ke-i dan peubah ke-k pada kelompok ke-j dan objek ke-i
- X_{ik} : input objek ke-i dan peubah ke-k
- Y_{jk} : output kelompok ke-j pada objek ke-k
- i : 1, 2, ..., n
- k : 1, 2, ..., p
- j : 1, 2, 3
- α_0 : *learning rate* (tingkat pembelajaran) awal

X dan Y dapat dinotasikan dalam bentuk matriks sebagai berikut:

$$X_{ik} = \begin{bmatrix} X_{11}, \dots, X_{1p} \\ X_{21}, \dots, X_{2p} \\ \vdots \\ X_{n1}, \dots, X_{np} \end{bmatrix} \quad Y_{ji} = \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ \vdots & \vdots & \vdots \\ Y_{n1} & Y_{n2} & Y_{n3} \end{bmatrix}$$

Kohonen (1989) menjelaskan bahwa setiap *neuron output* mempunyai bobot untuk masing-masing *neuron input*. Proses *learning* dilakukan dengan melakukan penyesuaian terhadap setiap bobot pada *neuron output*. Setiap *input* yang diberikan dihitung jarak Euclid-nya dengan setiap *neuron output*, kemudian dipilih *neuron output* yang mempunyai jarak minimum. *Neuron* yang mempunyai jarak yang paling kecil disebut *neuron terpilih* atau *neuron* yang paling sesuai dengan *input* yang diberikan. Jarak yang digunakan adalah jarak Euclid karena untuk mengukur kesamaan besarnya jarak suatu garis lurus yang menghubungkan antar objek. Jarak Euclid dapat dihitung dengan menggunakan rumus sebagai berikut:

$$d_i^2 = \|X_i - W_j\|^2 = \sum_{k=1}^p (X_{ik} - W_{jk})^2 \quad 2.4$$

dimana:

- d_i^2 : jarak *Euclid* Kuadrat objek pengamatan ke- i
- X_i : input objek pengamatan yang berukuran $i \times k$
- W_j : pembobot objek pengamatan yang berukuran $j \times k$
- n : banyaknya objek pengamatan
- i : $1, 2, \dots, n$
- k : $1, 2, \dots, p$
- j : $1, 2, 3$

Rumus 2.6 digunakan untuk menentukan *neuron* terpilih di mana jarak objek yang paling dekat dianggap sebagai *neuron* terpilih sebelumnya. Setelah mendapatkan *neuron winner* maka setelah itu *update* nilai bobot *neuron* terpilih dan tetangganya dengan menggunakan *Hebbian Learning Rule* dengan perhitungan sebagai berikut:

$$W_{jk}(t+1) = W_{jk}(t) + \alpha(t)h_{ij}(t) * [X_{ik}(t) - W_{jk}(t)] \quad 2.5$$

di mana:

- W_{jk} : bobot kelompok ke- j dan peubah ke- k
- $\alpha(t)$: *learning rate*
- t : $1, \dots, 100$
- $h_{ij}(t)$: fungsi tetangga ; $0 < \alpha(t) \leq \alpha(t-1) \leq 1$

Nilai awal untuk $W_{jk}(t)$ adalah ditentukan acak yang dinotasikan dalam bentuk matriks dengan ukuran yang sama dengan matriks peubah asal (X_{ik}). Fungsi tetangga (*neighbourhood*)

digunakan sebagai dasar untuk menentukan *neuron* terpilih. *Neuron* terpilih digunakan sebagai awal perhitungan ukuran kesamaan. Berikut adalah perhitungan untuk fungsi tetangga (*neighbourhood Kernel*):

$$h_{ij}(t) = \exp\left(-\frac{\|D(X_{ik}, Y_{jk})\|^2}{2\sigma^2(t)}\right) \quad 2.6$$

di mana:

- $D(X_{ik}, Y_{jk})$: kombinasi jarak objek pengamatan ke- i dan peubah ke- k pada kelompok ke- j
- $\sigma^2(t)$: cakupan tetangga (*neighbourhood*) pada iterasi ke- t
- i : 1, 2, ..., n
- k : 1, 2, ..., p
- j : 1, 2, 3

Nilai dari $\sigma^2(t)$ semakin menurun dengan bertambahnya tingkat iterasi. $\sigma^2(t)$ nilainya didasarkan pada lebar kelompok dalam *neuron hexagonal*/segi enam (Kohonen, 1989).

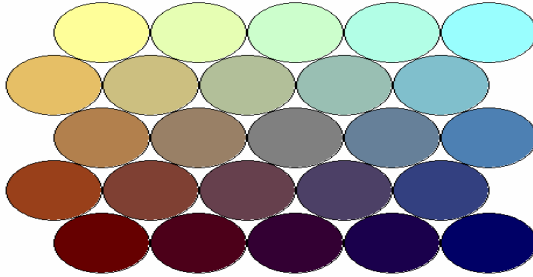
2.3.3 Visualisasi *Self Organizing Maps* (SOM)

Ciri dari SOM adalah untuk memvisualisasikan atau menggambarkan kelompok dari dimensi tinggi ke dalam dimensi dua. Dengan menggunakan SOM semua data yang digunakan dalam proses *training* akan direpresentasikan dalam bentuk *neuron* (Kohonen, 1995).

Tipe dari SOM dimensi dua adalah membentuk dan menggambarkan kelompok yang berdimensi dua. Selain itu, SOM memberikan informasi mengenai komposisi dari sampel dalam bentuk gambar dengan bentuk jaringan *neuron* dimensi dua. Dengan merepresentasikan gambar dimensi dua dari SOM, teknik visualisasi digunakan untuk menggambarkan ruang dimensi data dalam bentuk pengelompokan *neuron* (Ultsch, 1993).

Tipe dari visualisasi menggunakan Gambar 2.4 sebagai ilustrasi jarak antara jaringan pembobot. Perbedaan warna merepresentasikan lebar dari cakupan tetangga (*neighbourhood*).

Mapping plot



Gambar 2.4 Contoh dari representasi gambar SOM skala jarak 5x5 *Hexagonal*

Gambar 2.4 adalah ilustrasi hasil pengelompokan objek pengamatan dengan menggunakan SOM di mana objek yang mempunyai kesamaan warna adalah memiliki jarak paling dekat sehingga dimasukkan dalam satu kelompok. Perbedaan warna menunjukkan bahwa objek pengamatan terletak dalam kelompok yang berbeda.

2.3.4 Warna Dalam *Self Organizing Maps (SOM)*

Warna yang ditampilkan dalam metode SOM terdiri dari 3 komponen utama, yaitu RGB (*Red, Green, and Blue*). Warna yang ditampilkan dalam *layer hexagonal* 5x5 didasarkan pada jarak kedekatan antar *neuron (neighbourhood)*. Jarak antar *neuron* 5x5 *hexagonal* disajikan pada lampiran 2. Perubahan warna selama proses iterasi dipengaruhi oleh tingkat *learning*. *Layer hexagonal*/segi enam 5x5 digunakan karena mempunyai keuntungan dibandingkan *layer* lain yaitu jarak Euclid antar *neuron* yang berseberangan mempunyai jarak yang sama untuk semua enam *neuron* tetangga yang berada didekatnya dan *layer hexagonal* 5x5 mempunyai keuntungan yaitu tidak terdapat tumpang tindih antar objek yang masuk dalam satu kelompok terhadap kelompok lain (Vesanto, 2002).

Dalam *layer hexagonal* SOM selain berdasarkan perbedaan warna untuk membedakan antara kelompok satu dengan yang lain,

terdapat tanda segitiga, lingkaran, tanda plus atau yang lain untuk membedakan objek pengamatan mana saja yang termasuk dalam suatu kelompok. Warna yang dominan ditentukan oleh kadar/konsentrasi dari peubah yang menyusun suatu objek pengamatan (Kohonen, 1988).



Gambar 2.5 Skala Warna RGB (*Red, Green, dan Blue*)

Untuk warna merah bernilai 0 °, hijau bernilai 120 °, dan biru bernilai 240 ° dan untuk mengetahui nilai warna RGB dalam *Self Organizing Maps* (SOM) dapat dinotasikan sebagai berikut:

$$\tanh_{Pr_{eucilhexagon}} = \frac{\sqrt{3}(G - B)}{2.R - G - B} \quad 2.7$$

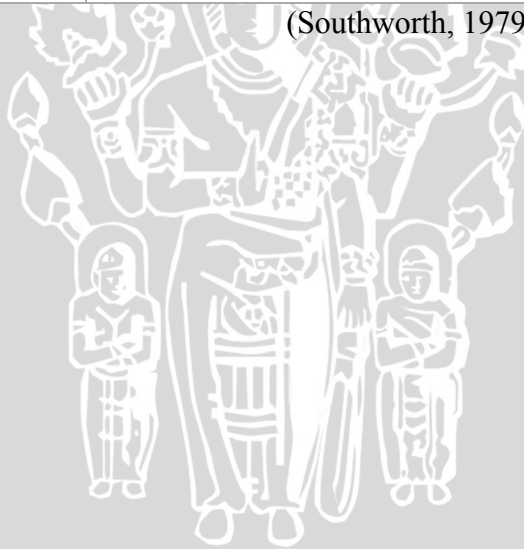
Pergeseran warna selama proses *Self Organizing Maps* (SOM) dapat terjadi karena perubahan tingkat *learning* (pembelajaran). Pergeseran warna dalam SOM mengikuti aturan RGB dan dapat diketahui nilainya berdasarkan rumus yang terdapat pada Tabel 2.1.

Tabel 2.1 Rumus Pergeseran Warna RGB

Urutan	Wilayah	Rumus
$R \geq G \geq B$	Merah-Kuning	$h_{Preucil\ circle} = 60^\circ \cdot \frac{G - B}{R - B}$
$G > R \geq B$	Kuning-Hijau	$h_{Preucil\ circle} = 60^\circ \cdot \left(1 + \frac{R - B}{G - B}\right)$

$G \geq B > R$	Hijau-Cyan	$h_{\text{Preucil circle}} = 60^\circ \cdot \left(2 + \frac{B - R}{G - R}\right)$
$B > G > R$	Cyan-Biru	$h_{\text{Preucil circle}} = 60^\circ \cdot \left(3 + \frac{G - R}{B - R}\right)$
$B > R \geq G$	Biru-Magenta	$h_{\text{Preucil circle}} = 60^\circ \cdot \left(4 + \frac{R - G}{B - G}\right)$
$R \geq B > G$	Magenta-Merah	$h_{\text{Preucil circle}} = 60^\circ \cdot \left(5 + \frac{B - G}{R - G}\right)$

(Southworth, 1979)



BAB III METODOLOGI

3.1 Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang merupakan hasil penelitian Forina *et al.* (2004) yang bersumber dari <http://archive.ics.uci.edu/ml/datasets/wine>, yang selanjutnya disebut Data *Wines* (data selengkapnya terdapat pada lampiran 1).

- **Data *Wines***

Data *wines* terbagi menjadi 177 objek pengamatan (jenis *wines*) dan masing-masing objek pengamatan terdiri atas 13 peubah, yaitu:

1. *Alcohol*
2. *Malic acid*
3. *Ash*
4. *Alcalinity Ash*
5. *Magnesium*
6. *Total Phenol*
7. *Flavonoids*
8. *Nonflavanoid Phenols*
9. *Proanthocyanins*
10. *Color Intensity*
11. *Hue*
12. *OD280/OD0315 of diluted wines*
13. *Proline*.

Peubah di atas adalah materi yang terkandung dalam *wines*. Software yang digunakan adalah program R.

Data yang digunakan dalam penelitian ini menggunakan data *wines*. Data *wines* merupakan data yang diperoleh dari hasil analisis kimia yang melibatkan proses fermentasi. Dalam proses fermentasi terdapat beberapa variabel yang digunakan sebagai dasar untuk mengetahui perkembangan dari proses fermentasi *wines*. Variabel tersebut adalah: *Alcohol*, *Malic acid*, *Ash*, *Alcalinity Ash*, *Magnesium*, *Total Phenol*, *Flavonoids*, *Nonflavanoid Phenols*, *Proanthocyanins*, *Color Intensity*, *Hue*, *OD280/OD0315*, *Proline*.

Alcohol adalah senyawa kimia yang digunakan secara luas dalam industri dan sains sebagai pereaksi, pelarut, dan bahan bakar.

Ada lagi *alcohol* yang digunakan secara bebas, yaitu yang dikenal sebagai spirtus. *Alcohol* digunakan secara bebas sebagai bahan makanan atau minuman, maka *alcohol* tersebut didenaturasi. Denaturasi *alcohol* disebut juga *methylated spirit*.

Malic acid disebut juga asam hijau apel. *Malic acid* terdapat dalam *wines*. *Malic acid* memberikan rasa panas dalam *wines*, walaupun jumlahnya menurun dengan meningkatnya kematangan *wines*. Proses fermentasi *malolactic* mengkonversi *malic acid* lebih banyak menjadi asam susu. *Malic acid* adalah sumber dari rasa kegetiran dan panas.

Ash atau soda *ash* adalah istilah komersial umum untuk *grade sodium carbonate*. Bahan ini digunakan sebagai *neutralizing agent* (penetralisir) dalam berbagai proses kimia.

Alkalinity Ash adalah tingkat kepekatan *grade sodium carbonate* dalam proses penetralisir suatu proses kimia.

Magnesium adalah unsur kimia penting yang terdapat pada *wines* yang berguna untuk membantu proses fermentasi.

Total Phenols adalah sejumlah senyawa yang mirip *alcohol* namun tidak termasuk *alcohol* karena mempunyai keunikan sifat pada rantai hidroksil. *Phenols* mempunyai berfungsi sebagai pemberi cita rasa pada *wines*.

Flavonoids adalah senyawa kimia yang terdapat pada *wines* yang berperan penting dalam menentukan metabolisme *wines* yaitu menentukan rasa dan warna pada *wines*.

Non Flavonoids Phenols adalah senyawa *phenolic* konsentrasi tertinggi ditemukan di dalam *wines* yang berfungsi untuk mengontrol pertumbuhan bakteri atau jamur pada saat proses fermentasi.

Proanthocyanins adalah senyawa kimia yang mempengaruhi perubahan warna selama proses fermentasi *wines*. Variasi warna pada *wines* merah sebagian berasal dari ionisasi dari *anthocyanin* Pigmen disebabkan oleh keasaman dari *wines*. Dalam hal ini, tiga jenis *anthocyanin* pigmen adalah merah, biru, dan warna lain dengan berbagai konsentrasi pigmen. *Wines* dengan pH rendah (tingkat keasaman tinggi) akan lebih tinggi terjadinya ionisasi yang akan

meningkatkan jumlah pigmen merah cerah. *Wines* dengan pH yang lebih tinggi akan memiliki konsentrasi yang lebih tinggi dan bergigmen biru.

Color Intensity adalah intensitas warna selama proses fermentasi dalam pembentukan *wines* berlangsung. Intensitas warna dapat diketahui perbedaannya pada saat akhir proses pembentukan *wines*.

Hue adalah salah satu jenis properti utama dari warna yang dapat dijelaskan dengan nama seperti merah, kuning, hijau, biru dan sebagainya.

$OD280/OD315$ adalah rasio perbandingan perubahan intensitas warna selama proses pembentukan *wines*.

Proline adalah senyawa kimia dalam *wines* yang berfungsi sebagai katalis (perantara) untuk membantu laju kecepatan proses kimiawi pada saat proses fermentasi *wines*.

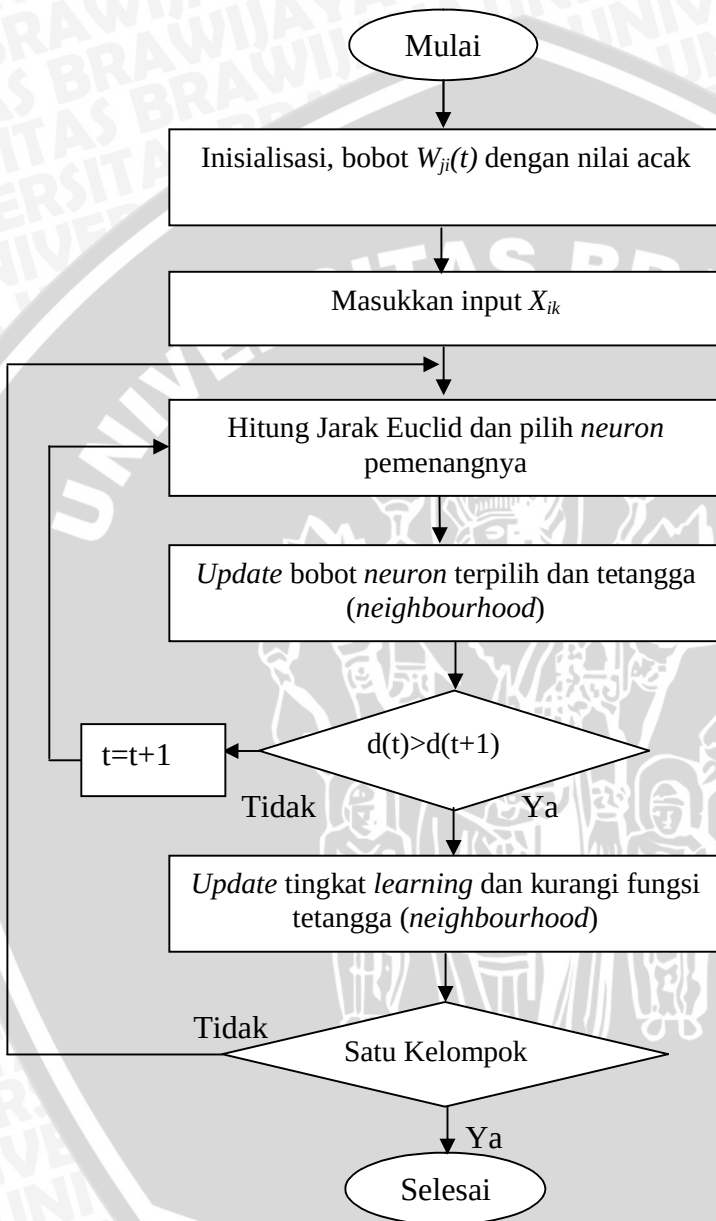
(Forina *et. al.*, 2004)

3.2 Metode

Metode yang dipakai dalam penelitian ini adalah Metode *Self Organizing Maps* (SOM).

Langkah langkah yang dilakukan pada data adalah sebagai berikut:

1. Inisialisasi bobot $W_{ji}(t)$ dengan nilai acak.
2. Masukan input X_{ik}
3. Hitung kedekatan jarak menggunakan jarak Euclid dan pilih *neuron* pemenang dengan persamaan 2.3 dan persamaan 2.4.
4. *Update* bobot *neuron* terpilih dan tetangganya (*neighbourhood*) dengan persamaan 2.5.
5. Jika jarak sebelumnya lebih kecil dari jarak setelah *update* maka jarak Euclid dan *update* nilai *neuron* dengan persamaan 2.4.
6. *Update* tingkat *learning* dan kurangi fungsi tetangga dengan persamaan 2.6
7. Jika belum terbentuk satu kelompok maka lakukan perhitungan jarak Euclid dan *update* nilai *neuron* dengan persamaan 2.3 dan persamaan 2.4
8. Lakukan sampai terbentuk satu kelompok.



Gambar 3.1 Diagram Alir Analisis Kelompok NonHirarki Self Organizing Maps (SOM).

BAB IV

HASIL DAN PEMBAHASAN

4.1. Analisis Kelompok NonHirarki

Dalam analisis kelompok nonhirarki jumlah kelompok yang dihasilkan ditentukan pada awal proses analisis. *Self Organizing Maps* (SOM) merupakan salah satu metode dalam analisis kelompok nonhirarki.

Dalam penelitian ini jumlah kelompok yang dibentuk adalah 3. Data yang digunakan dalam penelitian ini adalah data *wines* yang tersusun atas 3 tipe. Ketiga tipe data *wines* tersebut adalah *cultivars1* yang tersusun atas 3 peubah yang kadarnya tinggi yaitu *alcohol*, *malic acid* dan *total phenols*, *cultivars2* yang tersusun atas 3 peubah yang kadarnya tinggi yaitu *ash alkalinity*, *ash* dan *non flavonoids phenols*, dan *cultivars3* yang tersusun atas 3 peubah yang kadarnya tinggi yaitu *proline*, *proanthocyanins* dan *OD ratio*. Sedangkan jarak kedekatan antar objek menggunakan Jarak Euclid.

Matriks peubah asal $X_{i,k}$ dengan ukuran 177x3 merupakan inisialisasi awal untuk menghitung jarak Euclid dan matriks W_{jk} sebagai pembobot peubah dengan ukuran matriks 3x3. Nilai matriks W_{jk} didapat secara acak menggunakan tabel bilangan acak dan dinyatakan sebagai berikut:

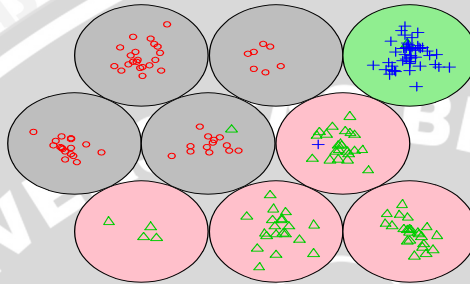
$$X_{ik} = \begin{bmatrix} 1 & 0 & 0 \\ \cdot & \cdot & \cdot \\ 0 & 1 & 0 \\ \cdot & \cdot & \cdot \\ 0 & 0 & 1 \end{bmatrix} \quad W_{jk} = \begin{bmatrix} 0.6 & 0.2 & 0.9 \\ 0.1 & 0.7 & 0.1 \\ 0.9 & 0.9 & 0.2 \end{bmatrix}$$

Untuk matriks X_{ik} dapat disederhanakan menjadi:

$$X_{ik} = \begin{bmatrix} 59 & 0 & 0 \\ 0 & 59 & 0 \\ 0 & 0 & 59 \end{bmatrix}$$

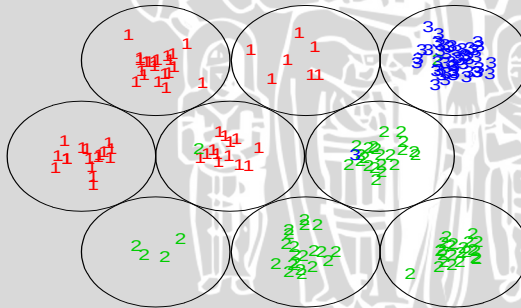
4.1.2 Visualisasi Data Wines dengan SOM

another mapping plot

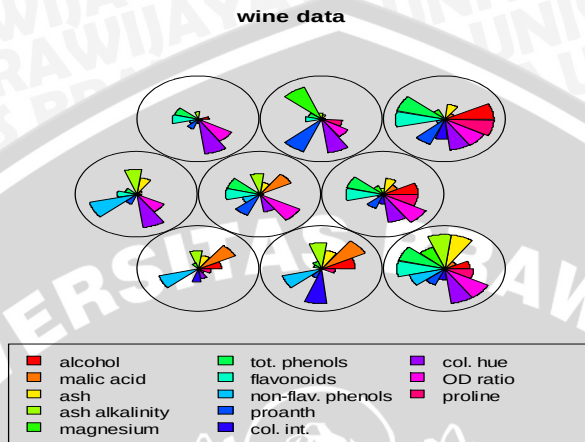


Gambar 4.1 Mapping Plot Data Wines 3x3 Layer Hexagonal.

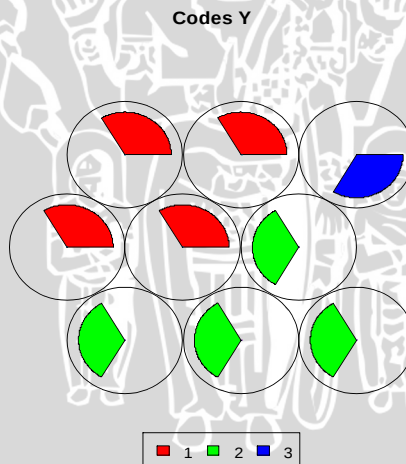
mapping plot



Gambar 4.2 Visualisasi SOM Distribusi Objek Pengamatan dalam Kelompok 3x3 Layer Hexagonal.

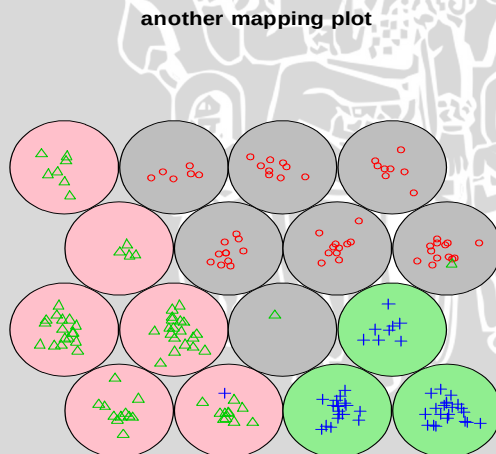


Gambar 4.3 Visualisasi SOM Konsentrasi Bahan Penyusun Wines
3x3 Layer Hexagonal.



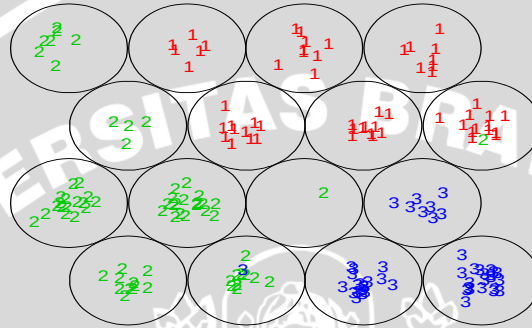
Gambar 4.4 Visualisasi SOM Kelompok yang Terbentuk 3x3 Layer
Hexagonal

Visualisasi SOM pada Gambar 4.1, Gambar 4.2 dan Gambar 4.4 menjelaskan bahwa terdapat tiga kelompok yang terbentuk. Tiap-tiap kelompok terdapat beberapa objek pengamatan yang mempunyai kedekatan jarak. Tanda segitiga, lingkaran dan plus adalah objek-objek pengamatan yang mempunyai kedekatan jarak yang menyusun kelompok terbentuk. Namun masih terdapat beberapa objek pengamatan yang masuk dalam kelompok yang bukan kelompok aslinya (tumpang tindih). Pada *layer hexagonal* baris kedua, *neuron* pada *neuron* kedua terdapat objek pengamatan yang seharusnya masuk dalam kelompok 2 namun masuk dalam kelompok 1 sedangkan pada *neuron* yang ketiga terdapat objek pengamatan yang masuk dalam kelompok 3 namun masuk dalam kelompok 2. Sehingga *layer hexagonal* 3x3 masih belum optimal karena terdapat beberapa objek pengamatan yang tidak masuk dalam kelompok aslinya. Pada Gambar 4.3 merupakan visualisasi dari kadar bahan penyusun *wines* dengan 13 peubah. Peubah yang mempunyai kadar yang tinggi digambarkan dengan sisi yang panjang sedangkan peubah dengan kadar rendah akan mempunyai sisi yang rendah.



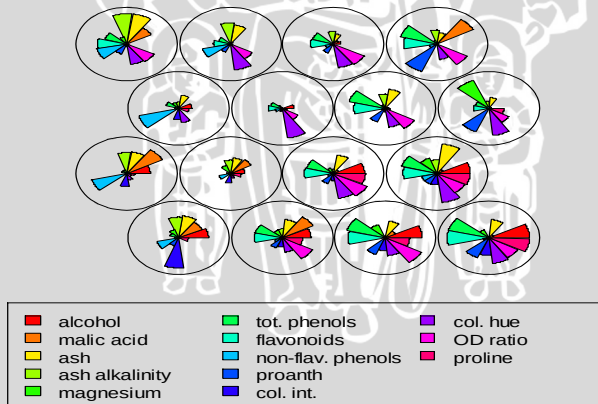
Gambar 4.5 Mapping Plot Data Wines 4x4 Layer Hexagonal.

mapping plot

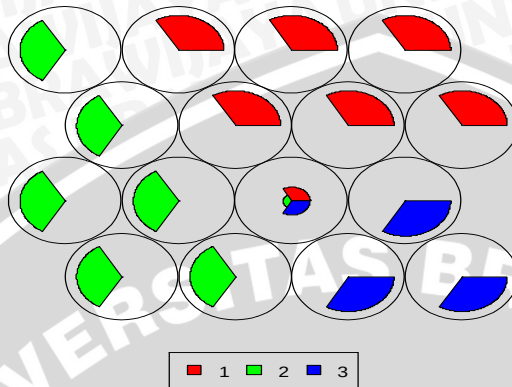


Gambar 4.6 Visualisasi SOM Distribusi Objek Pengamatan dalam Kelompok 4x4 Layer Hexagonal.

wine data

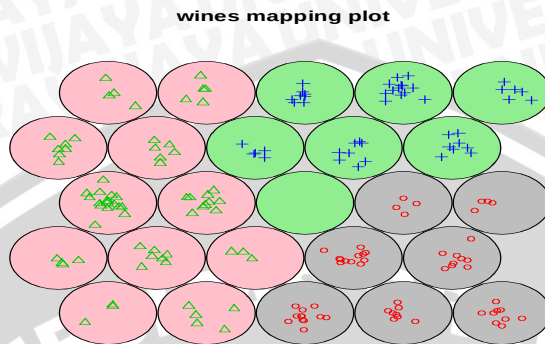


Gambar 4.7 Visualisasi SOM Konsentrasi Bahan Penyusun Wines 4x4 Layer Hexagonal.



Gambar 4.8 Visualisasi SOM Kelompok yang Terbentuk 4x4 *Layer Hexagonal*.

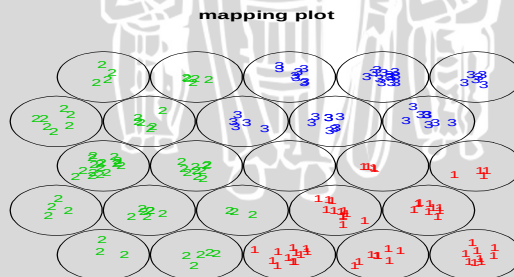
Visualisasi SOM pada Gambar 4.5, Gambar 4.6 dan Gambar 4.8 menjelaskan bahwa terdapat tiga kelompok yang terbentuk dengan 4x4 *layer hexagonal*. Tiap-tiap kelompok terdapat beberapa objek pengamatan yang mempunyai kedekatan jarak. Tanda segitiga, lingkaran dan plus adalah objek-objek pengamatan yang mempunyai kedekatan jarak yang menyusun kelompok terbentuk. Namun masih terdapat beberapa objek pengamatan yang masuk dalam kelompok yang bukan kelompok aslinya (tumpang tindih). Pada *layer hexagonal* baris kedua, *neuron* pada neuron keempat terdapat objek pengamatan yang seharusnya masuk dalam kelompok. Sehingga *layer hexagonal* 4x4 masih belum optimal karena terdapat objek pengamatan yang masuk dalam kelompok yang bukan kelompok aslinya. Pada Gambar 4.7 dan Gambar 4.8 merupakan visualisasi dari konsentrasi bahan penyusun *wines* dengan 13 peubah. Peubah yang mempunyai kadar yang tinggi digambarkan dengan sisi yang panjang sedangkan peubah dengan kadar rendah akan mempunyai sisi yang rendah.



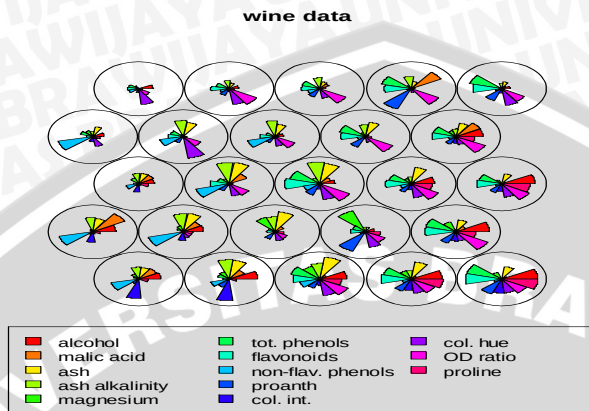
Gambar 4.9 *Mapping Plot Data Wines 5x5 Layer Hexagonal.*

Visualisasi SOM pada Gambar 4.9 dan Gambar 4.12 menjelaskan bahwa terdapat tiga kelompok yang terbentuk dengan 5x5 *layer hexagonal*. Tiap-tiap kelompok terdapat beberapa objek pengamatan yang mempunyai kedekatan jarak. Tanda segitiga, lingkaran dan plus adalah objek-objek pengamatan yang mempunyai kedekatan jarak yang menyusun kelompok terbentuk. Sehingga dari Gambar 4.9 tiga kelompok dapat diketahui pembagian objek-objek pengamatan yang masuk dalam tiap-tiap kelompok sudah merata dan tidak terdapat objek pengamatan yang masuk dalam kelompok yang lain.

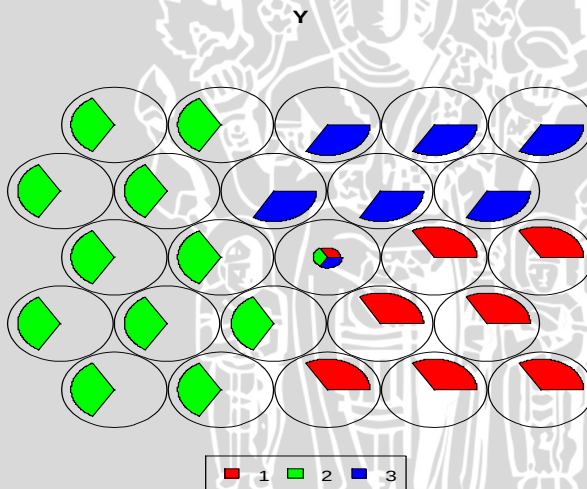
Pada Gambar 4.10 adalah distribusi objek pengamatan yang dalam ketiga kelompok yang terbentuk. Objek pengamatan yang mempunyai kedekatan jarak di inisialisasi dengan 1, 2 dan 3.



Gambar 4.10 Visualisasi SOM Distribusi Objek Pengamatan dalam Kelompok 5x5 *Layer Hexagonal*.



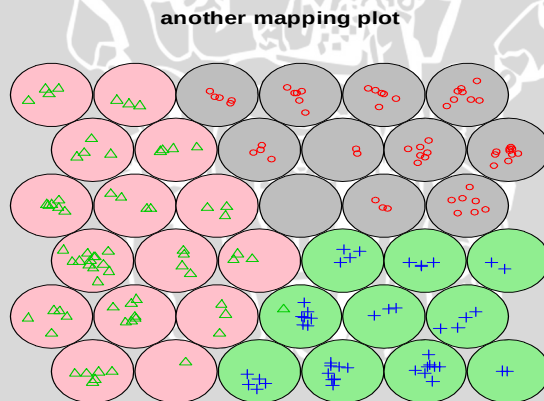
Gambar 4.11 Visualisasi SOM Konsentrasi Bahan Penyusun Wines
5x5 Layer Hexagonal.



Gambar 4.12 Visualisasi SOM Kelompok yang Terbentuk 5x5 Layer
Hexagonal.

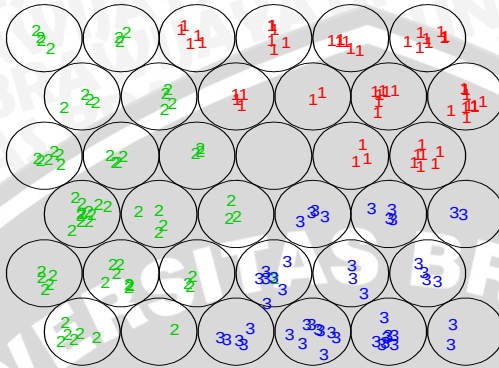
Pada Gambar 4.12 dan Gambar 4.16 adalah *output* yang terbentuk yaitu 3 kelompok. Masing-masing kelompok mempunyai warna yang berbeda. Warna tersebut mengikuti kaidah warna dalam SOM yaitu RGB (*Red*, *Green* dan *Blue*). Pada *neuron* di bagian tengah *layer* terdapat 3 warna sekaligus hal ini adalah karena tidak terdapat objek pengamatan yang masuk dalam *neuron* tersebut.

Dari Gambar 4.11 diketahui bahwa dalam *neuron* terdapat beberapa variabel pengamatan yang menyusun *wines*. Masing-masing variabel dalam *neuron* dibedakan menurut warna-warna yang berbeda dan ukuran panjang yang berbeda. Hal ini ditentukan oleh besar kecilnya konsentrasi peubah tersebut dalam membentuk *wines*. Semakin besar konsentrasi peubah yang menyusun objek pengamatan tersebut maka ukuran peubah tersebut dalam *neuron* semakin panjang. Semakin kecil konsentrasi peubah yang menyusun objek pengamatan tersebut maka ukuran peubah tersebut semakin pendek. Selain dapat menggambarkan konsentrasi pembentuk *wines*, SOM juga dapat menggambarkan pengelompokan data *wines* berdasarkan distribusi objek pengamatan.



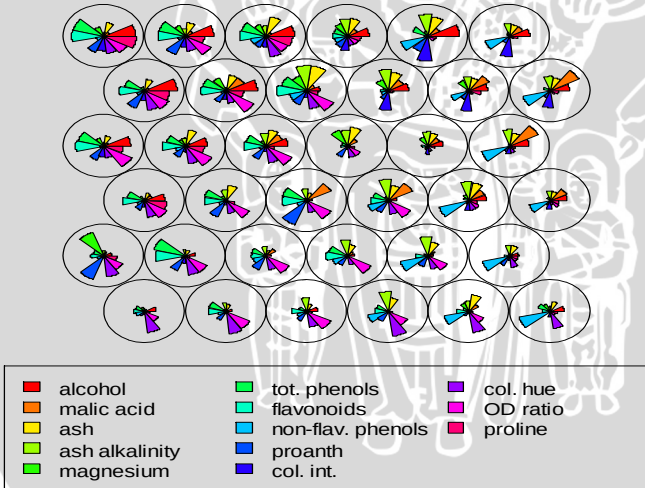
Gambar 4.13 Mapping Plot Data Wines 6x6 Layer Hexagonal.

mapping plot

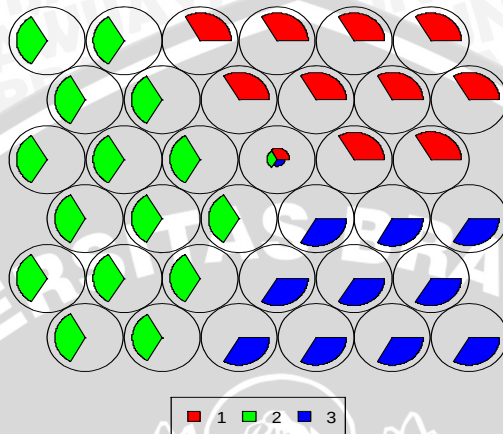


Gambar 4.14 Visualisasi SOM Distribusi Objek Pengamatan dalam Kelompok 6x6 *Layer Hexagonal*.

wine data



Gambar 4.15 Visualisasi SOM Konsentrasi Bahan Penyusun Wines 6x6 *Layer Hexagonal*.



Gambar 4.16 Visualisasi SOM Kelompok yang Terbentuk 6x6 Layer Hexagonal.

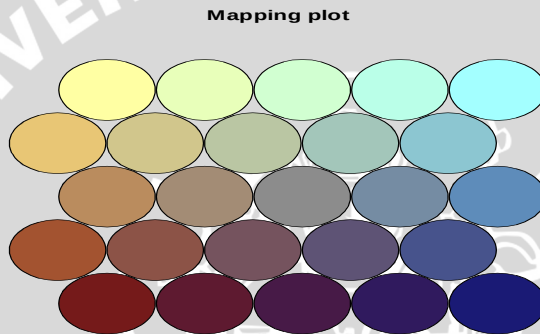
Dari Gambar 4.15 diketahui bahwa dalam *neuron* terdapat beberapa variabel pengamatan yang menyusun *wines*. Masing-masing variabel dalam *neuron* dibedakan menurut warna-warna yang berbeda dan ukuran panjang yang berbeda. Hal ini ditentukan oleh besar kecilnya konsentrasi peubah tersebut dalam membentuk *wines*. Semakin besar konsentrasi peubah yang menyusun objek pengamatan tersebut maka ukuran peubah tersebut dalam *neuron* semakin panjang. Semakin kecil konsentrasi peubah yang menyusun objek pengamatan tersebut maka ukuran peubah tersebut semakin pendek. Selain dapat menggambarkan konsentrasi pembentuk *wines*, SOM juga dapat menggambarkan pengelompokan data *wines* berdasarkan distribusi objek pengamatan.

Visualisasi SOM pada Gambar 4.13 dan Gambar 4.16 menjelaskan bahwa terdapat tiga kelompok yang terbentuk dengan 6x6 layer hexagonal. Tiap-tiap kelompok terdapat beberapa objek pengamatan yang mempunyai kedekatan jarak. Tanda segitiga, lingkaran dan plus adalah objek-objek pengamatan yang mempunyai kedekatan jarak yang menyusun kelompok terbentuk. Sehingga dari Gambar 4.9 tiga kelompok dapat diketahui pembagian objek-objek pengamatan yang masuk dalam tiap-tiap kelompok sudah merata dan

tidak terdapat objek pengamatan yang masuk dalam kelompok yang lain.

Dari Gambar 4.1 sampai 4.16 dapat diketahui bahwa *Layer Hexagonal* yang optimal dalam pengelompokan objek pengamatan adalah 5×5 *Layer Hexagonal* karena dalam *layer* tersebut objek-objek pengamatan sudah masuk dalam kelompok yang sesuai dan tidak ada objek pengamatan yang saling tumpang tindih.

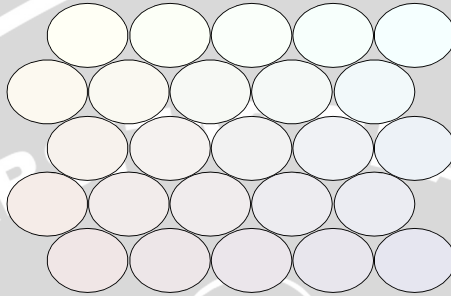
Pada Gambar 4.13 tingkat pembelajaran (*learning*) yang digunakan adalah 0.1 agar warna yang ditampilkan lebih jelas.



Gambar 4.17 Visualisasi SOM dengan Tingkat *Learning* 0.1 5×5 *Layer Hexagonal*.

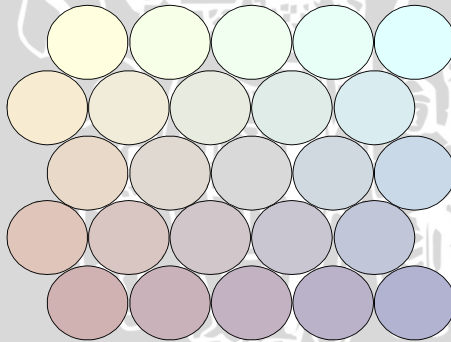
Visualisasi SOM dengan tingkat *learning* 0.1 merupakan yang paling jelas tingkat ketajaman warna. Kelompok yang terbentuk dapat diketahui perbedaannya dengan tingkat ketajaman warna yang berbeda pada tampilan Gambar 4.13. Berikut dibawah ini adalah tingkat perbandingan nilai *learning* terhadap perubahan ketajaman warna dari vsualisasi data wines dengan *Self Organizing Maps* (SOM):

Mapping plot



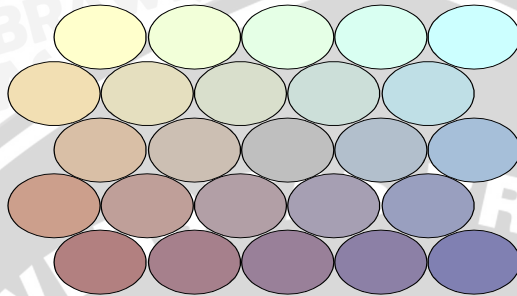
Gambar 4.18 Visualisasi SOM dengan Tingkat *Learning* 0.9 5x5 Layer Hexagonal.

Mapping plot



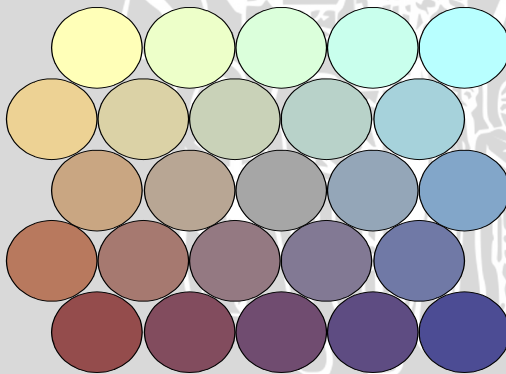
Gambar 4.20 Visualisasi SOM dengan Tingkat *Learning* 0.7 5x5 Layer Hexagonal.

Mapping plot



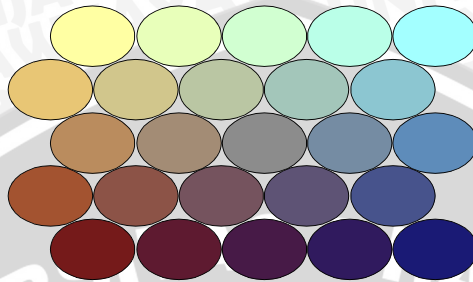
Gambar 4.19 Visualisasi SOM dengan Tingkat *Learning* 0.5 5x5
Layer Hexagonal.

Mapping plot



Gambar 4.21 Visualisasi SOM dengan Tingkat *Learning* 0.3 5x5
Layer Hexagonal.

Mapping plot



Gambar 4.22 Visualisasi SOM dengan Tingkat *Learning* 0.1 5x5 Layer Hexagonal.

Tabel 4.1 Hasil Pengelompokan Objek Pengamatan dengan SOM

Kelompok [1]												
176	71	21	13	43	137	59	166	29	78	169	39	128
16	74	14	91	2	157	51	101	47	155	140	152	10
95	163	145	54	100	170	27	149	55	121	177	111	117
64	110	52	103	171	142	173	11	107	116	124	73	161
97	168	89	48	20	23							

Kelompok [2]												
156	33	120	134	88	80	62	150	53	154	37	46	31
138	123	65	162	58	45	44	113	92	87	85	24	70
109	8	9	69	130	93	30	1	165	164	60	119	147
167	66	15	105	115	38	114	82	25	68	3	172	135
108	144	4	122	40	63	35	106	131	49	77	79	57
12	75	17	146	126	143							

Kelompok [3]												
6	41	36	67	139	90	28	76	118	125	102	72	19
7	133	50	151	83	42	160	158	94	32	5	127	61
18	153	84	159	96	129	141	98	175	81	174	104	99
112	136	56	26	34	148	132	22	86				

Pada Tabel 4.2 Kelompok 1 terdiri atas 58 objek pengamatan, kelompok 2 terdiri atas 71 objek pengamatan dan kelompok 3 terdiri atas 48 objek pengamatan.

Tabel 4.2 Nilai Tengah Peubah Wines Hasil *Training*

No	Peubah Wines	Nilai Tengah
1	<i>Alcohol</i>	12.9958333
2	<i>Malic Acid</i>	2.2941667
3	<i>Ash</i>	2.3580833
4	<i>Ash Alkalinity</i>	19.58983333
5	<i>Magnesium</i>	10.06166667
6	<i>Total Phenols</i>	2.2994167
7	<i>Flavonoids</i>	2.0177500
8	<i>Non Flavonoids Phenols</i>	0.3611667
9	<i>Proanthocyanins</i>	1.6215833
10	<i>Color Intensity</i>	5.0035833
11	<i>Color Hue</i>	0.9651333

No	Peubah <i>Wines</i>	Nilai Tengah
12	<i>OD Ratio</i>	2.6265000
13	<i>Proline</i>	0.7499750

Pada Tabel 4.3 nilai tengah (*centered*) peubah *wines* yang besar setelah *training* berturut-turut adalah *ash alkalinity*, *alcohol* dan *magnesium* sebesar 19.598333; 12.9953 dan 10.061. Hal ini mengindikasikan bahwa selama *training* peubah *ash alkalinity*, *alcohol* dan *magnesium* mempunyai kontribusi dalam pembentukan *wines*.

SOM merupakan teknik visualisasi dan reduksi dimensi pada proses pengelompokan objek pengamatan. Proses visualisasi data *wines* dilakukan dengan membentuk peubah awal dalam bentuk matriks peubah awal dengan pembobot untuk menghitung jarak kedekatan antar objek.

Berdasarkan Gambar 4.1 diketahui bahwa karakteristik yang membentuk kelompok 1 didasarkan pada konsentrasi bahan pembentuk data *wines* yang paling banyak yaitu *alcohol*, *malic acid* dan *total phenols*. Karakteristik yang membentuk kelompok 2 adalah *alkalinity*, *ash* dan *non flavonoids phenols*. Karakteristik yang membentuk kelompok 3 adalah *proline*, *proanthocyanins* dan *OD ratio*. Pada Gambar 4.1 belum diketahui dengan jelas mengenai batasan dari 3 kelompok yang terbentuk. Oleh karena itu, data *wines* diplotkan dengan Gambar 4.2.

Pada Gambar 4.2 terdapat tanda plus, segitiga dan lingkaran sebagai inisialisasi untuk objek pengamatan yang ada dalam satu kelompok. Objek pengamatan yang mempunyai tanda sama, maka objek pengamatan tersebut masuk dalam satu kelompok. Kelompok 1 (warna *pink*) simbol segitiga menjelaskan bahwa pembentukan kelompok didasarkan pada kadar *alcohol*, *ash* dan *proline*. Kelompok 2 (warna hijau) simbol tanda plus (+) pembentukan kelompoknya didasarkan pada kadar *OD280/OD0315 of diluted wines*, *pro anth* dan *malic acid*. Kelompok 3 (warna *silver*) tanda lingkaran pembentukan kelompoknya didasarkan pada kadar

flavonoids, *col. hue* dan *ash alkalinity*. Dengan menggunakan tingkat *learning* 0.1 maka data *wines* dapat divisualisasikan seperti pada Gambar 4.3. Warna yang ditampilkan pada gambar 4.3 didasarkan pada jarak kedekatan antara *neuron* dengan mengikuti aturan RGB. Nilai jarak kedekatan antar *neuron* (*neighbourhood*) disajikan pada lampiran 2.

Visualisasi jumlah *output* (kelompok yang terbentuk) yang disajikan pada Gambar 4.4. Gambar 4.4 dalam *layer hexagonal* terdapat angka 1, 2, dan 3. Angka tersebut merupakan inisialisasi dari kelompok yang terbentuk. Distribusi dari objek pengamatan yang termasuk dalam kelompok 1, kelompok 2, dan kelompok 3 disajikan dalam Tabel 4.2.



BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari penelitian ini dapat ditarik kesimpulan berikut:

1. Visualisasi data *wines* terdiri atas 3 kelompok. Kelompok 1 pengelompokannya didasarkan pada banyaknya kadar *alcohol*, *ash* dan *proline* terdiri atas 58 objek pengamatan, kelompok 2 pengelompokannya didasarkan pada *OD ratio*, *pro anth* dan *malic acid* terdiri atas 71 objek pengamatan, dan kelompok 3 pengelompokannya didasarkan pada banyaknya kadar *Col. Hue*, *flavonoids* dan *ash alkalinity* terdiri atas 48 kelompok.
2. Nilai *learning* dan *layer maps* yang optimal adalah nilai *learning* 0.1 dan *layer maps hexagonal* 5x5.

5.2 Saran

Disarankan pada peneliti lain untuk menggunakan metode *Self Organizing Maps* (SOM) yang diterapkan untuk dimensi 3 pada data *wines*.

UNIVERSITAS BRAWIJAYA



DAFTAR PUSTAKA

- Aas, K. 2001. **Microarray Data Mining: A Survey**.
<<http://www.nr.no/files/samba/smbi/microarraysurvey.pdf>>.
Tanggal akses: 9 Februari 2009.
- Chester, M. 1993. **Neural Network**. Prentice Hall. New Jersey.
- Duran, B. S. & Odell, P. L. 1974. **Cluster Analysis a Survey**. Springer-Verlag. New York.
- Dunn G, Everitt B. 1982. **An Introduction to Mathematical Taxonomy**. Cambridge University Press. Cambridge. MA.
- Forina et al. 2004. **An Extendible Package for Data Exploration, Classification and Correlation**. Institute of Pharmaceutical and Food Analysis and Technologies.
<http://archive.ics.uci.edu/ml/datasets/wine>. **Data Wines Set**.
Tanggal akses 15 Januari 2009.
- Forina et al. 2004. **An Extendible Package for Data Exploration, Classification and Correlation**. Institute of Pharmaceutical and Food Analysis and Technologies.
<http://id.wikipedia.org/wiki/Alkohol>
Tanggal akses 15 Januari 2009.
- Hartigan, J. A. 1967. **Clustering Algoritms**. John Wiley and Sons. New York.
- Johnson, R.A. dan D.W. Wincern. 1992. **Applied Multivariate Statistical Analysis**. Prentice Hall Inc. New Jersey.
- Kohonen. T. 1988. **Self Organizing Maps and Associate Memory**. Springer-Verlag. New York.
- Kohonen. T. 1989. **Self Organization and Associative Memory**. Springer-Verlag. New York.
- Kohonen, T. 1995. **Self Organizing Maps**. Springer-Verlag. Berlin.
- Kohonen, T. 2001. **Self Organizing Maps**. Springer-Verlag. Berlin.

- Lance, G. N. & Williams, W. T. 1966. **A Generalized Sorting Strategy for Computer Classification**. Springer-Verlag. New York.
- Limin, F. 1994. **Neural Network in Computer Intelligence**. McGraw Hill. New York.
- Matjijik, A.A., I.M. Sumertajaya, H. Wijayanto. 2002. **Aplikasi Analisis Peubah Ganda**. Jurusan Statistika FMIPA IPB. Bogor.
- Sharma, S. 1996. **Applied Multivariate Techniques**. John Wiley and Sons. New York.
- Sofan, L. and Guègan, J.F. 1999. **Artificial Neural Network as a Tool in Ecological Modeling**. McGraw Hill. New York.
- Southworth, Miles. 1979. **Color Separation Techniques**. Arts Design Livonia, New York.
- Tokukata, H. 1999. **Application of Self Organizing Maps to Chemical Data Analysis**. Springer-Verlag. Berlin.
- Ultsch, A. 1993. **Self Organized Feature Maps for Monitoring and Knowledge Acquisition of a Chemical Process**. Springer-Verlag. Berlin.
- Ultsch, A. 2007. **Emergence in Self Organized Feature Maps**. Springer-Verlag. Berlin.
- Vesanto, J. 2002. **Data Exploration Process Based on the Self Organizing Maps**. Springer-Verlag. Berlin.
- William, W.T. Lambert, J. M. & Lance, G. N. 1966. **Multivariate Methods in Plants Ecology**. Springer-Verlag. New York.

UNIVERSITAS BRAWIJAYA



Lampiran 1. Data Hasil Fermentasi *Wines* (dalam mg/L)

no	alcohol	malic acid	ash	ash alk.	magnesium	tot. phen.	flav.	non-flav.phen.	proant.	col.int.	hue	OD	proline
1	0.26	1.28	4.38	1.05	3.4	1.05	13.2	1.78	2.14	11.2	1	2.65	2.76
2	0.3	2.81	5.68	1.03	3.17	1.185	13.2	2.36	2.67	18.6	1.01	2.8	3.24
3	0.24	2.18	7.8	0.86	3.45	1.48	14.4	1.95	2.5	16.8	1.13	3.85	3.49
4	0.39	1.82	4.32	1.04	2.93	0.735	13.2	2.59	2.87	21	1.18	2.8	2.69
5	0.34	1.97	6.75	1.05	2.85	1.45	14.2	1.76	2.45	15.2	1.12	3.27	3.39
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
173	0.52	1.06	7.7	0.64	1.74	0.74	13.7	5.65	2.45	20.5	0.95	1.68	0.61
174	0.43	1.41	7.3	0.7	1.56	0.75	13.4	3.91	2.48	23	1.02	1.8	0.75
175	0.43	1.35	10.2	0.59	1.56	0.835	13.3	4.28	2.26	20	1.2	1.59	0.69
176	0.53	1.46	9.3	0.6	1.62	0.84	13.2	2.59	2.37	20	1.2	1.65	0.68
177	0.56	1.35	9.2	0.61	1.6	0.56	14.1	4.1	2.74	24.5	0.96	2.05	0.76

UNIVERSITAS BRAWIJAYA



Lampiran 2. Matriks Jarak Antar Neuron Dalam 5x5 Hexagonal SOM

	[,1]	[,2]	[,3]	[,4]	...	[,23]	[,24]	[,25]
[1,]	0.0000000	1.0000000	2.0000000	2.0000000	...	2.1794495	2.1794495	1.3228757
[2,]	1.0000000	0.0000000	1.0000000	2.0000000	...	1.3228757	2.1794495	2.1794495
[3,]	2.0000000	1.0000000	0.0000000	1.0000000	...	0.8660254	1.3228757	2.1794495
[4,]	2.0000000	2.0000000	1.0000000	0.0000000	...	1.3228757	0.8660254	1.3228757
[5,]	1.0000000	2.0000000	2.0000000	1.0000000	...	2.1794495	1.3228757	0.8660254
[6,]	1.0000000	1.7320508	2.6457513	1.7320508	...	3.0413813	2.2912878	1.8027756
...	.	.	.	.	.	.	.	.
...	.	.	.	.	.	.	.	.
...	.	.	.	.	.	.	.	.
...	.	.	.	.	.	.	.	.
[21,]	0.8660254	1.3228757	2.1794495	2.1794495	...	2.0000000	2.0000000	1.0000000
[22,]	1.3228757	0.8660254	1.3228757	2.1794495	...	1.0000000	2.0000000	2.0000000
[23,]	2.1794495	1.3228757	0.8660254	1.3228757	...	0.0000000	1.0000000	2.0000000
[24,]	2.1794495	2.1794495	1.3228757	0.8660254	...	1.0000000	0.0000000	1.0000000
[25,]	1.3228757	2.1794495	2.1794495	1.3228757	...	2.0000000	1.0000000	0.0000000

Lampiran 3. Indeks Pengelompokan Objek Pengamatan dengan SOM

[1]	176	71	21	13	43	137	59	166	29	78	169	39	128	16	74	14	91	2
[19]	157	51	101	47	155	140	152	10	95	163	145	54	100	170	27	149	55	121
[37]	177	111	117	64	110	52	103	171	142	173	11	107	116	124	73	161	97	168
[55]	89	48	20	23	156	33	120	134	88	80	62	150	53	154	37	46	31	138
[73]	123	65	162	58	45	44	113	92	87	85	24	70	109	8	9	69	130	93
[91]	30	1	165	164	60	119	147	167	66	15	105	115	38	114	82	25	68	3
[109]	172	135	108	144	4	122	40	63	35	106	131	49	77	79	57	12	75	17
[127]	146	126	143	6	41	36	67	139	90	28	76	118	125	102	72	19	7	133
[145]	50	151	83	42	160	158	94	32	5	127	61	18	153	84	159	96	129	141
[163]	98	175	81	174	104	99	112	136	56	26	34	148	132	22	86			

[illegible][illegible][illegible][illegible][illegible][illegible]

UNIVERSITAS BRAWIJAYA



Lampiran 5. Hasil *Training* dengan SOM

no	alcohol	malic acid	ash	ash alk.	magnesium	tot. phen.	flav.	non-flav.phen.	proant.	col.int.	hue	OD	proline
1	0.2112	0.257055	0.04	0.11629	1.2929621	-1.0121	-1.31	1.34573187	-0.272	1.893	-1.6	-1.4	2.88E-01
2	1.0479	-0.68138	1.09	1.56389	-0.975002	1.01391	0.825	-1.20491491	0.418	-0.716	1.7	0.76	-1.09E+00
3	-0.08	1.308445	1.02	-0.289	0.092275	0.17234	0.384	-0.88608406	0.603	-0.222	0.3	1.27	6.41E-02
4	2.1272	-0.49021	0.11	-2.3736	-0.641478	1.24768	1.638	0.54865475	2	0.175	1.2	0.15	1.28E+00
5	0.2961	1.464851	-0.24	-0.6075	0.1589799	0.53079	0.6	-0.32813008	0.065	-0.284	-0.6	0.53	-2.24E-01
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
172	0.49	-0.50914	0.92	-1.0242	-0.464764	0.8903	0.918	-0.17902815	-0.24	-0.11	-0.2	0.87	1.4287925
173	0.6384	-0.48234	1.03	-0.155	0.7346138	0.09217	0.507	-0.58014185	-0.082	-0.368	0.6	0.38	1.1112153
174	0.4035	0.804165	0.05	0.59443	-0.535315	-0.5783	-1.27	0.703422	-0.59	1.448	-1.8	-1.4	-0.302003
175	-0.227	-0.0267	0.12	1.34381	-0.112006	-1.8233	-0.93	-0.74058734	-1.324	0.278	-1.3	-1.8	-0.587823
176	0.8857	-0.42873	-0.02	-0.8744	0.0996492	0.5072	0.858	-0.74058734	0.18	-0.54	0.7	1.98	0.9206689
177	-1.031	-0.65208	-0.2	0.98411	-0.676419	-0.8177	-0.33	0.54297652	-0.047	-1.121	1.6	-0.5	-0.794248

Lampiran 6. Matriks awal peubah asal $X_{i,k}$

	1	2	3
[1,]	1	0	0
[2,]	1	0	0
[3,]	1	0	0
[4,]	1	0	0
[5,]	1	0	0
[...]	.	.	.
[...]	.	.	.
[60]	0	1	0
[...]	.	.	.
[...]	.	.	.
[...]	.	.	.
[117]	0	0	1
[...]	.	.	.
[...]	.	.	.
[...]	.	.	.
[...]	.	.	.
[...]	.	.	.
[177]	0	0	1



Lampiran 7. Macro R Data *Wines*

- **Menentukan matriks peubah asal $X_{i,k}$**

```
classes <- c(rep(1, 59), rep(2, 59), rep(3, 59))
```

```
classmat <- classvec2classmat(classes)
```

```
classmat
```

```
classmat2classvec(classmat)
```

- **Memvisualisasikan plot data wines layer 3x3**

```
data(wines)
```

```
wines.sc <- scale(wines)
```

```
set.seed(7)
```

```
wine.som <- som( data=wines.sc, grid = somgrid(3, 3, "hexagonal"))
```

```
plot(wine.som, main = "wine data")
```

- **Memvisualisasikan plot data wines layer 4x4**

```
data(wines)
```

```
wines.sc <- scale(wines)
```

```
set.seed(7)
```

```
wine.som <- som( data=wines.sc, grid = somgrid(4,4, "hexagonal"))
```

```
plot(wine.som, main = "wine data")
```

- **Memvisualisasikan plot data wines layer 5x5**

```
data(wines)
```

```
set.seed(7)
```

```

kohmap <- xyf(scale(wines), classvec2classmat(wine.classes),
grid = somgrid(5, 5, "hexagonal"), rlen=100)

plot(kohmap, type="changes")

plot(kohmap, type="codes", main = c(" X", "Y"))

plot(kohmap, type="counts")

plot(kohmap, type="quality")

plot(kohmap, type="mapping",
labels=wine.classes, col=wine.classes+1,
main="mapping plot")

xyfpredictions <-
classmat2classvec(predict(kohmap)$unit.predictions)

bgcols <- c("gray", "pink", "lightgreen")

plot(kohmap, type="mapping", col=wine.classes+1,
pchs=wine.classes, bgcol=bgcols[as.integer(xyfpredictions)],
main=" mapping plot")

```

- Memvisualisasikan plot data wines layer 6x6

```

data(wines)

wines.sc <- scale(wines)

set.seed(7)

wine.som <- som( data=wines.sc, grid = somgrid(6, 6, "hexagonal"))

plot(wine.som, main = "wine data")

```

- Visualisasi data wines dengan nilai *Learning* 0.9

```
data(wines)
```

```
som.wines <- som(wines, grid = somgrid(5, 5, "hexagonal"))
```

```
colour1 <- tricolor(som.wines$grid)
```

```
plot(som.wines, "mapping", bg = rgb(colour1))
```

```
colour2 <- tricolor(som.wines$grid, phi = c(pi/6, 0, -pi/6))
```

```
plot(som.wines, "mapping", bg = rgb(colour2))
```

```
colour3 <- tricolor(som.wines$grid, phi = c(pi/6, 0, -pi/6), offset = .9)
```

```
plot(som.wines, "mapping", bg = rgb(colour3))
```

- Visualisasi data wines dengan nilai *Learning* 0.7

```
data(wines)
```

```
som.wines <- som(wines, grid = somgrid(5, 5, "hexagonal"))
```

```
colour1 <- tricolor(som.wines$grid)
```

```
plot(som.wines, "mapping", bg = rgb(colour1))
```

```
colour2 <- tricolor(som.wines$grid, phi = c(pi/6, 0, -pi/6))
```

```
plot(som.wines, "mapping", bg = rgb(colour2))
```

```
colour3 <- tricolor(som.wines$grid, phi = c(pi/6, 0, -pi/6), offset = .7)
```

```
plot(som.wines, "mapping", bg = rgb(colour3))
```

- Visualisasi data wines dengan nilai *Learning* 0.5

```
data(wines)
```

```
som.wines <- som(wines, grid = somgrid(5, 5, "hexagonal"))
```

```
colour1 <- tricolor(som.wines$grid)
```

```
plot(som.wines, "mapping", bg = rgb(colour1))
```

```
colour2 <- tricolor(som.wines$grid, phi = c(pi/6, 0, -pi/6))
```

```
plot(som.wines, "mapping", bg = rgb(colour2))
```

```
colour3 <- tricolor(som.wines$grid, phi = c(pi/6, 0, -pi/6), offset = .5)
```

```
plot(som.wines, "mapping", bg = rgb(colour3))
```

- Visualisasi data wines dengan nilai *Learning* 0.3

```
data(wines)
```

```
som.wines <- som(wines, grid = somgrid(5, 5, "hexagonal"))
```

```
colour1 <- tricolor(som.wines$grid)
```

```
plot(som.wines, "mapping", bg = rgb(colour1))
```

```
colour2 <- tricolor(som.wines$grid, phi = c(pi/6, 0, -pi/6))
```

```
plot(som.wines, "mapping", bg = rgb(colour2))
```

```
colour3 <- tricolor(som.wines$grid, phi = c(pi/6, 0, -pi/6), offset = .3)
```

```
plot(som.wines, "mapping", bg = rgb(colour3))
```

- Visualisasi data wines dengan nilai *Learning* 0.1

```
data(wines)
```

```
som.wines <- som(wines, grid = somgrid(5, 5, "hexagonal"))
```

```
colour1 <- tricolor(som.wines$grid)
```

```
plot(som.wines, "mapping", bg = rgb(colour1))
```

```
colour2 <- tricolor(som.wines$grid, phi = c(pi/6, 0, -pi/6))
```

```
plot(som.wines, "mapping", bg = rgb(colour2))
colour3 <- tricolor(som.wines$grid, phi = c(pi/6, 0, -pi/6), offset = .1)
plot(som.wines, "mapping", bg = rgb(colour3))
```

- Matriks Jarak Antar Neuron

```
data(wines)
kohmap <- som(wines, grid = somgrid(5, 5, "hexagonal"), rlen=10)
par(mfrow=c(1,2))
dists <- unit.distances(kohmap$grid, toroidal=FALSE)
plot(kohmap, type="property", property=dists[1,],
     main="Distances to unit 1", zlim=c(0,6),
     palette = rainbow, ncolors = 7, contin = TRUE)
dists <- unit.distances(kohmap$grid, toroidal=TRUE)
plot(kohmap, type="property", property=dists[1,],
     main="Distances to unit 1 (toroidal)", zlim=c(0,6),
     palette = rainbow, ncolors = 7, contin = TRUE)
```

- Menentukan kesimpulan

```
data(wines)
xyf.wines <- xyf(scale(wines), classvec2classmat(wine.classes),
grid = somgrid(5, 5, "hexagonal"))
xyf.wines
summary(xyf.wines)
```