

## **BAB II**

### **KAJIAN TEORI**

Bab II berisi tentang kajian teori. Teori-teori yang digunakan dalam penelitian ini diantaranya mengenai kanker payudara, penelitian-penelitian terdahulu, pengolahan citra digital dan ekstraksinya, *Neural Network*, metode *Clustering*, dan hasil ketepatan klasifikasi.

#### **A. Kanker Payudara (*Breast Cancer*)**

Berikut adalah hal-hal mengenai Kanker Payudara:

##### **1. Pengertian Kanker Payudara**

Kanker payudara adalah tumor ganas yang dimulai di sel-sel payudara. Sebuah tumor ganas adalah sekelompok sel-sel kanker yang dapat tumbuh (menyerang) jaringan sekitarnya atau menyebar (Metastasis) ke daerah tubuh. Penyakit ini umumnya terjadi pada wanita, tetapi pria dapat pula terserang (*American Cancer Society, 2016*)

##### **2. Gejala Kanker Payudara**

Pada tahap awal kanker payudara, biasanya tidak dirasakan sakit atau bahkan sama sekali tidak ada gejala. Ketika kanker mulai berkembang, gejala berikut ini akan terlihat (*Yayasan Pita Merah Muda Bali, 2013*):

- a. Suatu gumpalan padat atau benjolan pada payudara atau sekitar ketiak.
- b. Puting susu yang tertarik kedalam
- c. Perubahan warna kulit payudara, pembengkakan atau jerawat (berkerut, berlesung ataupun pengelupasan).

- d. Darah atau cairan lain yang keluar dari puting susu.
- e. Perubahan ukuran atau bentuk payudara.

### **3. Penyebab Kanker Payudara**

Hal-hal yang menyebabkan terjadinya kanker payudara tidak diketahui secara pasti. Berkaitan dengan hal tersebut para ahli memiliki pendapat yang berbeda-beda. Begitu juga dengan faktor-faktor yang dapat mempengaruhi seseorang berpotensi terkena kanker payudara. Berikut beberapa faktor risiko kanker payudara menurut beberapa ahli :

a. Menurut (*Cancer Treatment Centers of America, 2015*):

1) Usia

Rata-rata, wanita berusia lebih dari 60 tahun lebih mungkin didiagnosis kanker payudara. Hanya sekitar 10 - 15 persen kanker payudara terjadi pada wanita berusia kurang dari 45 tahun. Namun, diagnosa ini mungkin berbeda untuk ras atau etnis yang berbeda.

2) Jenis kelamin

Meskipun hampir 2.000 orang didiagnosis kanker payudara setiap tahun, kanker payudara 100 kali lebih sering terjadi pada wanita. *The National Cancer Institute* memperkirakan bahwa lebih dari 190.000 perempuan akan didiagnosis terserang kanker payudara setiap tahunnya.

3) Riwayat keluarga

Memiliki riwayat keluarga kanker payudara, khususnya perempuan dengan ibu, saudara perempuan atau anak perempuan yang memiliki atau menderita kanker payudara, mungkin dua kali lipat berisiko.

4) Faktor mewarisi

Beberapa mutasi genetik yang diwariskan dapat meningkatkan risiko kanker payudara. Mutasi pada gen BRCA1 dan BRCA2 adalah penyebab paling umum yang diwariskan. mutasi langka lainnya juga dapat membuat beberapa wanita lebih rentan untuk terserang kanker payudara. Pengujian gen mengungkapkan adanya masalah genetik yang potensial, terutama di keluarga yang memiliki riwayat kanker payudara.

5) Obesitas

Setelah menopause, jaringan lemak dapat menyebabkan peningkatan kadar estrogen, dan tingkat estrogen yang tinggi yang dapat meningkatkan risiko kanker payudara. Berat badan saat dewasa dan kelebihan lemak tubuh di sekitar pinggang juga dapat menjadi penyebabnya.

6) Tidak memiliki anak

Wanita yang tidak memiliki anak, atau yang hamil di kemudian hari (di atas usia 35) mungkin memiliki risiko lebih besar terkena kanker payudara. Menyusui dapat membantu menurunkan risiko kanker payudara.

7) Kepadatan payudara tinggi

Wanita dengan jaringan lemak, jaringan kelenjar dan fibrosa sedikit lebih berisiko untuk terkena kanker payudara dibandingkan wanita dengan payudara kurang padat.

8) Perubahan payudara tertentu

kondisi payudara tertentu dapat meningkatkan risiko kanker payudara.

9) Sejarah menstruasi

Wanita yang mulai menstruasi pada usia dini (sebelum usia 12) dan/atau menopause pada usia yang lebih tua (setelah usia 55) memiliki risiko sedikit lebih tinggi terkena kanker payudara. Peningkatan risiko mungkin karena masa hidup yang lebih lama mengeluarkan hormon estrogen dan progesteron.

10) Gaya hidup

Aktivitas fisik dalam bentuk latihan rutin selama empat sampai tujuh jam seminggu dapat membantu mengurangi risiko kanker payudara.

11) Mengonsumsi Alkohol

Penggunaan alkohol terkait dengan meningkatnya risiko terserang kanker payudara. Risiko meningkat dengan jumlah alkohol yang dikonsumsi.

12) Pil KB

Menggunakan kontrasepsi oral dalam 10 tahun terakhir dapat meningkatkan risiko terkena kanker payudara. Risiko menurun dari waktu ke waktu setelah pil dihentikan.

13) Terapi hormon pascamenopause (PHT) yang dikombinasi

Menggunakan terapi hormon gabungan setelah menopause meningkatkan risiko terkena kanker payudara.

14) Paparan dietilstilbestrol (DES)

Penggunaan DES sebelumnya, obat yang biasa diberikan kepada wanita hamil (pada tahun 1940-1971) untuk mencegah keguguran, dapat sedikit meningkatkan risiko terkena kanker payudara. Wanita yang ibunya mengambil DES selama kehamilan juga mungkin memiliki risiko sedikit lebih tinggi terserang kanker payudara.

15) Paparan radiasi

Anak-anak atau remaja, yang sedang melakukan terapi radiasi pada daerah dada sebagai pengobatan kanker lain berisiko terserang kanker payudara.

b. Menurut (*American Cancer Society*, 2016):

1) Pernah menderita kanker payudara

Seorang wanita dengan kanker pada satu payudara memiliki peningkatan risiko terkena kanker baru di payudara lainnya atau di bagian lain dari payudara yang sama. (Hal ini berbeda dengan

kekambuhan dari kanker pertama.) Risiko ini bahkan lebih tinggi jika kanker payudara didiagnosis pada usia yang lebih muda.

2) Ras dan etnis

Secara keseluruhan, wanita kulit putih memiliki risiko terkena kanker payudara yang lebih kecil dari wanita Afrika-Amerika, tetapi wanita Afrika-Amerika lebih mungkin untuk meninggal karena kanker ini. Namun, pada wanita di bawah usia 45 tahun, kanker payudara lebih sering terjadi pada wanita Afrika-Amerika. Wanita Asia, Hispanik, dan penduduk asli Amerika memiliki risiko lebih rendah terkena kanker payudara dan kematian yang diakibatkannya.

3) Diet

Sebuah penelitian terbaru menemukan risiko lebih tinggi terkena kanker payudara pada wanita yang makan lebih banyak daging merah.

4) Asap tembakau

Dalam beberapa tahun terakhir, lebih banyak studi telah menemukan bahwa merokok dalam jangka panjang terkait dengan lebih tingginya risiko terkena kanker payudara. Beberapa studi telah menemukan bahwa risiko tertinggi pada kelompok tertentu, seperti wanita yang mulai merokok sebelum mereka memiliki anak pertama mereka.

5) Faktor Kontroversial dan tidak dibenarkan. Hal ini bisa berupa *antiperspirant*, penggunaan bra, aborsi, implan payudara, dan lain-lain.

#### **4. Klasifikasi Kanker Payudara**

Perkembangan kondisi abnormal payudara hingga menjadi sel kanker terbagi menjadi tiga kelas yaitu normal, tumor (*benign*) dan kanker (*malignant*) (*National Breast Cancer Foundation, 2015*). Berikut adalah penjelasan masing-masing klasifikasi kanker payudara:

a. **Diagnosis normal**

Payudara normal merupakan payudara dengan pertumbuhan sel normal, dimana sel-sel payudara yang tumbuh sama dengan sel-sel payudara yang rusak atau mati.

b. **Diagnosis tumor (*benign*)**

Tumor merupakan pertumbuhan sel yang abnormal dimana pembelahan sel pada payudara lebih cepat dari pada sel yang rusak atau mati. Meskipun tumor ini umumnya tidak membahayakan jaringan di sekitarnya, namun tumor dapat terus tumbuh dan menekan organ yang menyebabkan rasa sakit. Dalam situasi ini, tumor dapat diangkat untuk menghilangkan rasa sakit.

c. **Diagnosis kanker (*malignant*)**

Kanker merupakan tumor yang menjadi ganas. Tumor yang tumbuh menyerang dan merusak jaringan disekitarnya.

#### **5. Cara Pendeteksian Kanker Payudara**

Kanker payudara kadang-kadang ditemukan setelah gejala muncul, tetapi banyak wanita dengan kanker payudara dini tidak memiliki gejala. Inilah sebabnya mengapa mendapatkan tes skrining yang dianjurkan sebelum gejala

berkembang sangat penting. Tes pencitraan menggunakan x-ray, medan magnet, gelombang suara, atau zat radioaktif dapat membuat gambar dari bagian dalam tubuh. Tes pencitraan dapat dilakukan untuk sejumlah alasan, termasuk untuk membantu mengetahui area mencurigakan yang memungkinkan menjadi kanker, mengetahui seberapa jauh kanker mungkin telah menyebar, dan untuk membantu menentukan apakah pengobatan bekerja. Tes pencitraan ini yaitu (*American Cancer Society*, 2016):

a. *Mammogram*

*Mammogram* adalah x-ray untuk payudara. Skrining *mammogram* digunakan untuk mencari penyakit pada payudara wanita yang tidak memiliki tanda-tanda atau gejala dari masalah payudara. *Mammogram* biasanya mengambil 2 tampilan (gambar x-ray diambil dari sudut yang berbeda) dari masing-masing payudara. Untuk *mammogram*, payudara ditekan antara 2 lempengan untuk meratakan dan menyebarkan jaringan. Jika kanker ditemukan, diagnostik *mammogram* sering dilakukan untuk mendapatkan pandangan yang lebih menyeluruh dari kedua payudara. Hal ini untuk memeriksa setiap daerah abnormal lain yang bisa jadi merupakan kanker.

b. USG payudara

USG, juga dikenal sebagai sonografi, menggunakan gelombang suara untuk menguraikan bagian tubuh. Untuk tes ini, instrumen seperti mikrofon kecil disebut transduser ditempatkan pada kulit (yang sering dilumasi pertama kali dengan gel USG). Transduser memancarkan



gelombang suara dan mengambil gema saat dipantulkan jaringan tubuh. Gema dikonversi oleh komputer menjadi gambar hitam dan putih yang ditampilkan di layar komputer. Tes ini tidak menimbulkan rasa sakit dan tidak memaparkan radiasi. USG telah menjadi alat yang digunakan bersama dengan *mammogram* karena banyak tersedia dan lebih murah daripada yang lainnya, seperti MRI. Biasanya, USG payudara digunakan untuk menargetkan area spesifik yang ditemukan pada *mammogram*. USG membantu membedakan antara tumor jinak dan kanker. Dalam seseorang dengan tumor payudara, USG dapat juga digunakan untuk melihat pembesaran kelenjar getah bening di bawah lengan. Penggunaan USG untuk skrining kanker payudara tidak dianjurkan. Namun, kini uji klinis telah menemukan manfaat dari penggunaan USG setelah dilakukan skrining payudara dengan *mammogram* pada wanita dengan payudara padat yang berisiko lebih tinggi terkena kanker payudara.

c. *Magnetic resonance imaging (MRI)*

Scan MRI menggunakan gelombang radio dan magnet kuat dan bukan sinar-x. Energi dari gelombang radio yang diserap dan kemudian dirilis dalam pola yang dibentuk oleh jaringan tubuh dan oleh penyakit tertentu. Sebuah komputer menerjemahkan pola menjadi gambar yang sangat rinci. Dalam mencari kanker, cairan kontras disebut gadolinium disuntikkan ke pembuluh darah sebelum atau selama *scan* untuk menunjukkan detail yang lebih baik. *Scan* MRI dapat memakan waktu yang lama – hampir satu jam. Dalam pemeriksaan MRI, pasien harus

berbaring di dalam sebuah tabung sempit, menghadap ke bawah pada *platform*. *Platform* ini memiliki bukaan untuk setiap payudara yang memungkinkan untuk mengambil gambarnya tanpa kompresi. *Platform* ini berisi sensor yang diperlukan untuk menangkap gambar MRI. MRI dapat digunakan bersama dengan *mammogram* untuk skrining wanita yang berisiko tinggi terkena kanker payudara, atau dapat digunakan untuk mengkaji secara lanjut daerah yang mencurigakan yang sebelumnya telah ditemukan menggunakan *mammogram*. MRI juga sering digunakan untuk wanita yang telah didiagnosis kanker payudara dalam menentukan ukuran kanker yang sebenarnya dan untuk mencari kanker lainnya yang muncul di payudara.

## B. Pengolahan Citra Digital

Citra digital adalah representasi dari citra dua dimensi sebagai sebuah kumpulan nilai digital yang disebut elemen gambar atau piksel. Piksel adalah istilah yang sering digunakan untuk menunjukkan unsur-unsur pada citra digital tersebut (Gonzales & Woods, 2008:1). Citra dua dimensi adalah citra (gambar) yang terdiri dari panjang dan lebar.

Citra digital dapat ditulis dalam bentuk matrik sebagai berikut.

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, N - 1) \\ f(1,0) & f(1,1) & \dots & f(1, N - 1) \\ \vdots & \vdots & & \vdots \\ f(M - 1,0) & f(M - 1,1) & \dots & f(M - 1, N - 1) \end{bmatrix} \quad (2.1)$$

Suatu piksel memiliki nilai dalam rentang tertentu, dari nilai minimum sampai nilai maksimum. Jangkauan yang digunakan berbeda-beda tergantung dari

jenis warnanya. Namun secara umum jangkauannya adalah 0-255. Citra dengan penggambaran seperti ini digolongkan ke dalam citra integer (Darma Putra, 2010: 20). Citra digital terbagi menjadi beberapa jenis. Jenis citra digital yang sering digunakan adalah citra biner, citra grayscale, dan citra warna. Adapun penjelasan mengenai jenis citra adalah sebagai berikut (Usman Ahmad, 2005 : 14):

### **1. Citra biner**

Citra biner adalah citra digital yang memiliki dua kemungkinan nilai piksel yaitu hitam yang dinyatakan dengan 0 dan putih yang dinyatakan dengan 1. Nilai 0 (hitam) adalah *background points*, biasanya bukan merupakan bagian dari citra sesungguhnya. Sedangkan nilai 1 (putih) adalah *region points*, yaitu bagian dari citra sebenarnya (bukan latar belakang). Citra biner juga disebut B&W (*black and white*) atau monokrom. Proses pemineranan dilakukan dengan membulatkan ke atas atau ke bawah untuk setiap nilai keabuan dari piksel yang berada di atas atau di bawah nilai batas.

### **2. Citra grayscale**

Citra *grayscale* atau abu – abu merupakan citra yang mampu menghasilkan gradasi warna abu-abu dari warna hitam hingga warna putih. Tingkat keabuan disini merupakan warna abu dengan berbagai tingkatan, berupa bilangan antara 0 s.d 255. Nilai ini digunakan untuk menunjukkan nilai intensitas. Nilai 0 untuk warna hitam, nilai 255 untuk warna putih dan nilai antara 0 sd 255 untuk warna antara hitam dan putih (keabuan).

### 3. Citra warna

Citra warna atau yang biasa disebut dengan citra RGB adalah citra digital yang setiap pikselnya mewakili warna kombinasi dari tiga warna dasar yaitu R = *Red*, G = *Green*, dan B = *Blue*. Citra warna disebut juga *true color* karena mempunyai jumlah warna yang cukup besar yaitu mencapai 16 juta warna.

#### C. Ekstraksi Citra

Proses ekstraksi merupakan salah satu karakteristik penting yang digunakan dalam mengidentifikasi objek atau pola citra, karena metode ekstraksi citra yang tepat akan mampu memberikan informasi yang detail tentang kelas suatu citra. Pada penelitian ini, metode ekstraksi citra yang digunakan adalah metode *Gray Level Co-occurrence Matrix* (GLCM). GLCM merupakan suatu metode ekstraksi citra yang banyak digunakan dalam klasifikasi citra dan merupakan salah satu metode yang cukup efektif dalam melakukan klasifikasi karena mampu memberikan informasi yang detail tentang suatu citra dalam hal tekstur (Gadkari, 2004). GLCM adalah suatu matriks yang elemen-elemennya merupakan jumlah pasangan piksel yang memiliki tingkat kecerahan tertentu (Toni, 2013: 18).

Ekstraksi citra yang dilakukan dengan metode GLCM dapat menghasilkan 14 fitur ekstraksi. Ekstraksi citra dapat dilakukan dengan bantuan aplikasi MATLAB R2013a dengan *script* MATLAB R2013a yang terlampir pada Lampiran 3.

Fitur – fitur ekstraksi yang didapatkan adalah *energy*, *contrast*, *correlation*, *sum of square variance*, *Inverse Difference Moment (IDM)*, *sum average*, *sum variance*, *sum entropy*, *entropy*, *difference variance*, *difference entropy*, *maximum probability*, *homogeneity* dan *dissimilarity*.

### 1. *Entropy* (Entropi)

Nilai entropi menunjukkan keteracakan distribusi derajat keabuan suatu citra. Semakin acak distribusi derajat keabuannya, semakin tinggi nilai entropi yang dihasilkan (Yegar & Chairisni, 2013: 6). Rumus *entropy* adalah sebagai berikut (Mohanaiah, *et al.*, 2013: 2):

$$Entropy = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} -P_{ij} \cdot \log P_{ij} \quad (2.2)$$

Dimana,

$i, j$  = koordinat spasial fungsi  $P_{ij}$

$P_{ij}$  = peluang nilai level keabuan pada baris ke-  $i$  dan kolom ke-  $j$

$N_g$  = level abu-abu

### 2. *Energy* (Energi)

Nilai energi bertolak belakang dengan entropi. Semakin tinggi nilai entropi maka nilai energi akan semakin rendah. Hal ini dikarenakan, nilai energi menggambarkan keteraturan penyebaran derajat suatu citra keabuan (Yegar & Chairisni, 2013). Rumus *energy* adalah sebagai berikut (Girisha *et al.*, 2013):

$$Energy = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P_{ij}^2 \quad (2.3)$$

$i, j$  = koordinat spasial fungsi  $P_{ij}$

$P_{ij}$  = peluang nilai level keabuan pada baris ke-  $i$  dan kolom ke-  $j$

$N_g$  = level abu-abu

### 3. **Contrast (kontras)**

Fitur kontras digunakan untuk menghitung *range* perbedaan derajat keabuan dalam sebuah citra. Semakin jauh perbedaan derajat keabuan setiap pasangan piksel, semakin tinggi nilai kontras. Demikian sebaliknya, jika perbedaan derajat keabuan setiap pasangan piksel tidak signifikan, nilai kontras akan rendah (Yegar & Chairisni, 2013). Rumus *contrast* adalah sebagai berikut (Girisha *et al*, 2013):

$$contrast = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 \cdot P_{ij} \quad (2.4)$$

$i, j$  = koordinat spasial fungsi  $P_{ij}$

$P_{ij}$  = peluang nilai level keabuan pada baris ke-  $i$  dan kolom ke-  $j$

$N_g$  = level abu-abu

### 4. **Correlation (korelasi)**

Korelasi adalah fitur yang digunakan untuk menghitung ketergantungan linear sebuah citra. Jika derajat keabuan antar pasangan piksel memiliki hubungan linear maka nilai korelasi akan menjadi tinggi (Yegar & Chairisni, 2013: 6). Rumus *correlation* adalah sebagai berikut (Girisha *et al*, 2013: 2719):

$$correlation = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i \cdot j \cdot P_{ij}) - \mu_x \cdot \mu_y}{\sigma_x \cdot \sigma_y} \quad (2.5)$$

dengan,

$i, j$  = koordinat spasial fungsi  $P_{ij}$

$P_{ij}$  = peluang nilai level keabuan pada baris ke-  $i$  dan kolom ke-  $j$

$N_g$  = level abu-abu

Dimana,  $\mu_x, \mu_y, \sigma_x$  dan  $\sigma_y$  adalah rata-rata dan standar deviasi dari peluang marginal  $P_x(i)$  dan  $P_y(j)$  yang diperoleh dari menjumlahkan baris dan kolom dari masing-masing matriks  $P_{ij}$

### 5. *Sum of Squares (Variance)*

*Variance* adalah ukuran heterogenitas atau variasi elemen-elemen matriks. *Variance* meningkat ketika nilai tingkat keabuan berbeda dari rata-ratanya dengan kata lain, citra dengan tingkat keabuan kecil maka akan memiliki *Variance* yang kecil pula (Gadkari, 2004: 13). Rumus *Sum of Square (Variance)* adalah sebagai berikut (Haralick, Shanmugam & Dinstein, 1973: 619):

$$Variance = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu)^2 p(i, j) \quad (2.6)$$

$i, j$  = koordinat spasial fungsi  $P_{ij}$

$P_{ij}$  = peluang nilai level keabuan pada baris ke-  $i$  dan kolom ke-  $j$

$N_g$  = level abu-abu

### 6. *Inverse Difference Moment (IDM)*

IDM adalah ukuran dari homogenitas lokal. Nilai IDM tinggi ketika level abu-abu (*gray level*) lokal seragam dan invers dari GLCM tinggi. Rumus IDM adalah sebagai berikut (Mohanaiah *et al*, 2013: 2):

$$IDM = \frac{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P_{ij}}{1+(i-j)^2} \quad (2.7)$$

$i, j$  = koordinat spasial fungsi  $P_{ij}$

$P_{ij}$  = peluang nilai level keabuan pada baris ke-  $i$  dan kolom ke-  $j$

$N_g$  = level abu-abu

### 7. *Sum Average*

*Sum Average* adalah fitur yang menunjukkan seberapa banyak nilai rata-rata piksel yang ada dalam citra. Rumus *Sum Average* (SA) adalah sebagai berikut (Haralick, Shanmugam & Dinstein, 1973: 619):

$$\text{Sum Average} = \sum_{k=2}^{2N_g} k P_{x+y}(k) \quad (2.8)$$

Dengan,

$$P_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j); i + j = k; k = 2, 3, \dots, 2N_g$$

$N_g$  = level abu-abu

### 8. *Sum Entropy*

*Sum Entropy* adalah fitur yang menunjukkan seberapa banyak level keabu – abuan yang acak. Rumus *Sum Entropy* (SE) adalah sebagai berikut (Haralick, Shanmugam & Dinstein, 1973: 619):

$$\text{Sum Entropy} = - \sum_{k=2}^{2N_g} P_{x+y}(k) \log\{P_{x+y}(k)\} \quad (2.9)$$

Dengan,

$$P_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j); i + j = k; k = 2, 3, \dots, 2N_g$$

$N_g$  = level abu-abu

### 9. *Sum Variance*

*Sum Variance* menyatakan heterogenitas spasial (perbedaan) gambar. Rumus (Abouelatta, 2013: 217). *Sum Variance* (SV) adalah sebagai berikut (Haralick, Shanmugam & Dinstein, 1973: 619):



$$\text{Sum Variance} = \sum_{i=2}^{2N_g} (i - SE)^2 P_{x+y}(k) \quad (2.10)$$

Dengan,

$$P_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j); i + j = k; k = 2, 3, \dots, 2N_g$$

$N_g$  = level abu-abu

### 10. *Difference Variance*

*Difference Variance* (DV) menyatakan ukuran variabilitas lokal (Abouelatta, 2013: 217). Rumus *Difference Variance* adalah sebagai berikut (Haralick, Shanmugam & Dinstein, 1973: 619):

$$DV = \text{variansi dari } P_{x+y}(k) \quad (2.11)$$

Dengan,

$$P_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j); i + j = k; k = 2, 3, \dots, 2N_g$$

### 11. *Difference Entropy*

*Difference Entropy* (DE) adalah ukuran variabilitas perbedaan mikro (lokal) (Abouelatta, 2013: 217). Rumus *Difference Entropy* adalah sebagai berikut (Haralick, Shanmugam & Dinstein, 1973:619):

$$DE = - \sum_{i=0}^{N_g-1} P_{x-y}(k) \log\{P_{x+y}(k)\} \quad (2.12)$$

Dengan,

$$P_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j); i + j = k; k = 2, 3, \dots, 2N_g$$

$N_g$  = level abu-abu

## 12. Maximum Probability

*Maximum probability* menghitung tingkat keabu-abuan, yang mempunyai peluang maksimum (*maximum probability*) pada GLCM.

$$\text{Maximum Probability} = \max_{i,j}\{p(i,j)\} \quad (2.13)$$

$i, j$  = koordinat spasial fungsi  $P_{ij}$

$P_{ij}$  = peluang nilai level keabuan pada baris ke-  $i$  dan kolom ke-  $j$

## 13. Homogeneity (Homogenitas)

Fitur homogenitas akan menghitung keseragaman variasi derajat keabuan sebuah citra. Fitur homogenitas akan memiliki nilai yang tinggi derajat keabuan yang hampir sama (Yegar & Chairisni, 2013: 6).

Rumus homogenitas adalah sebagai berikut (Girisha *et al*, 2013: 2720):

$$\text{Homogeneity} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1+(i-j)^2} \cdot P_{ij} \quad (2.14)$$

Dengan,

$i, j$  = koordinat spasial fungsi  $P_{ij}$

$P_{ij}$  = peluang nilai level keabuan pada baris ke-  $i$  dan kolom ke-  $j$

$N_g$  = level abu-abu

## 14. Dissimilarity

*Dissimilarity* mirip dengan kontras (*contrast*). *Dissimilarity* akan tinggi ketika daerah lokalnya memiliki kontras yang tinggi. Fitur ini sensitif terhadap variabilitas spasial tingkat abu-abu dan input gambar (Abouelatta, 2013: 216):

$$\text{Dissimilarity} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P(i,j) |i-j| \quad (2.15)$$

Dengan,

$i, j$  = koordinat spasial fungsi  $P_{ij}$

$P_{ij}$  = peluang nilai level keabuan pada baris ke-  $i$  dan kolom ke-  $j$

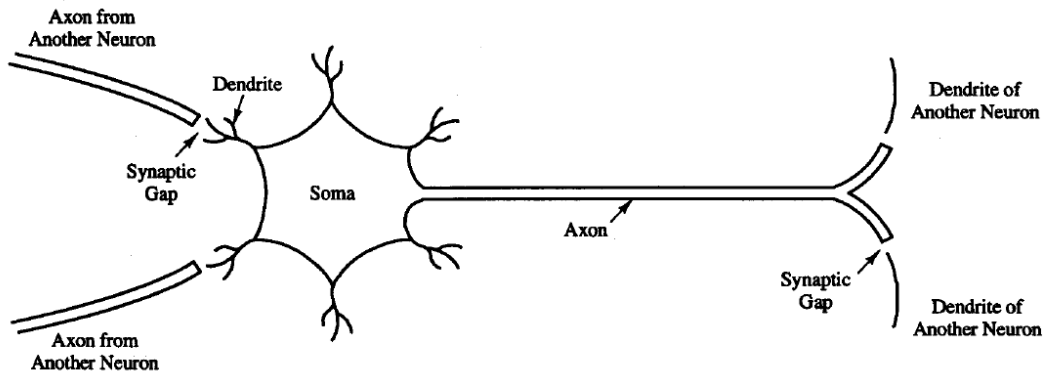
$N_g$  = level abu-abu

#### D. Neural Network (NN)

*Artificial Neural Network* (ANN) atau yang biasa disebut *Neural Network* (NN) merupakan sistem pemrosesan informasi yang memiliki karakteristik mirip dengan jaringan saraf biologis (Fausett, 1994: 3). *Neural Network* dibentuk sebagai generalisasi model matematika dari jaringan saraf biologi (Jong Jek Siang, 2005: 2). Manusia memiliki sekitar  $10^{11}$  *neuron* yang terus menerus membenahi diri atau direorganisasi oleh pengaruh eksternal (sekitar  $10^5$  *neuron* dapat hancur selama keadaan mabuk, beberapa jenis makanan atau pengaruh lingkungan juga dapat merusak sel-sel otak) (Kriesel, 2005: 4-5).

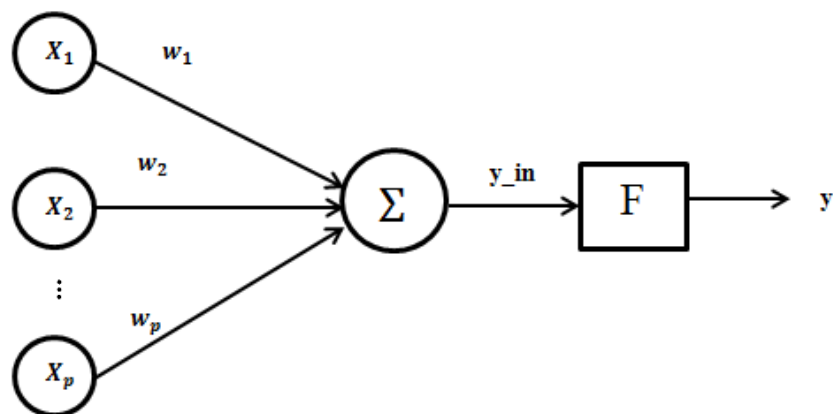
Suatu *Neural Network* ditandai dengan (1) arsitektur yaitu pola dari hubungan antar *neuron*, (2) algoritma pembelajaran yaitu metode untuk menentukan bobot pada hubungan *neuron* dan (3) fungsi aktivasi (Fausett, 1994: 3). Sebuah *neuron* memiliki tiga bagian utama yaitu sel tubuh atau soma (dimana nukleus/inti sel terletak), dendrit, dan akson (Gambar 2.1). Dendrit adalah sambungan tunggal berbentuk silinder yang memanjang dari sel tubuh dan membawa impuls (sinyal) dari *neuron* (Lin & Lee, 1995: 206). Sinyal tersebut berupa impuls elektrik yang dikirim melalui celah sinaptik melalui proses kimiawi. Sinyal yang masuk dimodifikasi (diperkuat/ diperlemah). Selanjutnya,

soma atau sel tubuh menjumlahkan semua sinyal yang masuk. Apabila sinyal yang diterima cukup kuat, maka sinyal akan diteruskan ke sel lain melalui akson (Fausett, 1994: 5).



**Gambar 2.1.** Syaraf Secara Biologi

Seperti halnya otak manusia, *Neural Network* juga terdiri dari beberapa *neuron*, dan ada hubungan antara *neuron-neuron* tersebut. *Neuron-neuron* tersebut akan mentransformasikan informasi yang diterima melalui sambungan keluarannya menuju ke *neuron-neuron* yang lain. Pada *Neural Network*, hubungan ini dikenal dengan nama bobot. Informasi tersebut disimpan pada suatu nilai tertentu pada bobot tersebut. Gambar 2.2 menunjukkan struktur *neuron* pada jaringan syaraf.



**Gambar 2.2.** *Neural Network* Sederhana

Dalam *Neural Network*, informasi yang disebut dengan *input* akan dikirim ke *neuron* dengan bobot tertentu. *Input* ini akan diproses oleh suatu fungsi perambatan yang akan menjumlahkan nilai-nilai semua bobot. Hasil penjumlahan ini kemudian akan dibandingkan dengan suatu nilai ambang (*threshold*) tertentu melalui fungsi aktivasi setiap *neuron*. Apabila *input* tersebut melewati suatu ambang tertentu, maka *neuron* tersebut akan diaktifkan, tetapi jika tidak, maka *neuron* tersebut tidak akan diaktifkan. Apabila *neuron* tersebut diaktifkan, maka *neuron* tersebut akan mengirimkan output melalui bobot-bobot outputnya ke semua *neuron* yang berhubungan dengannya, dan demikian seterusnya.

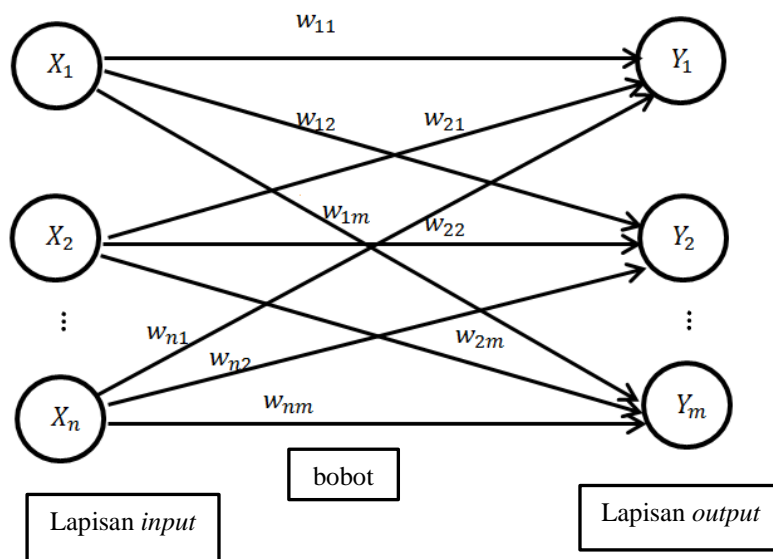
Pada *Neural Network*, *neuron-neuron* akan dikumpulkan dalam lapisan-lapisan (*layer*) yang disebut dengan lapisan *neuron* (*neuron layers*). *Neuron-neuron* pada satu lapisan akan dihubungkan dengan lapisan-lapisan sebelum dan sesudahnya (kecuali lapisan *input* dan *output*). Informasi yang diberikan pada *Neural Network* akan dirambatkan lapisan ke lapisan, mulai dari lapisan *input* sampai ke lapisan *output* melalui lapisan yang lainnya, yang disebut lapisan tersembunyi (*hidden layer*). Arah perambatan ini sesuai dengan algoritma pembelajaran yang digunakan (Sri Kusumadewi & Sri Hartati, 2010: 69-72). Menurut Fausett (1994:3) karakteristik dari *neural network* ditentukan oleh beberapa hal, yaitu: arsitektur, fungsi aktivasi, dan *learning algorithm* (algoritma pembelajaran).

## 1. Arsitektur

Hubungan antar *neuron* dalam *Neural Network* mengikuti pola tertentu tergantung pada arsitektur jaringan syarafnya. Menurut Fausett (1994: 12-15) terdapat 3 arsitektur dalam *Neural Network*, antara lain:

### a. Jaringan Layar Tunggal (*single layer network*)

Dalam jaringan ini, sekumpulan *input neuron* dihubungkan langsung dengan sekumpulan *outputnya*. Dalam beberapa model (misal *perceptron*), hanya ada sebuah *neuron output*.

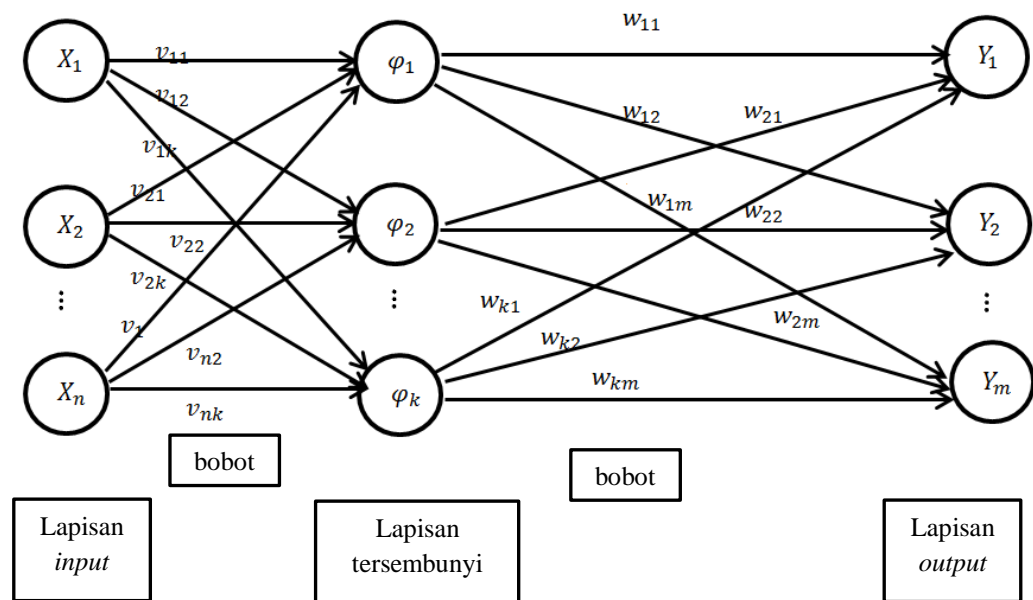


**Gambar 2.3.** Jaringan Layar Tunggal

Gambar 2.3 menunjukkan arsitektur jaringan dengan  $i$  *neuron input* ( $x_1, x_2, \dots, x_n$ ) dan  $s$  *neuron output* ( $Y_1, Y_2, \dots, Y_m$ ). Dalam jaringan ini, semua *neuron input* dihubungkan dengan semua *neuron output*, meskipun dengan bobot yang berbeda-beda. Tidak ada *neuron input* yang dihubungkan dengan *neuron input* lainnya. Demikian pula dengan *neuron output*.

b. Jaringan Layer Jamak (*multilayer network*)

Jaringan layer jamak merupakan jaringan dengan satu layer simpul atau lebih (disebut *hidden neuron/ neuron tersembunyi*) antara *neuron input* dan *neuron output*. Terdapat layer bobot antara dua tingkat *neuron* yang berdekatan (*input, hidden, output*).

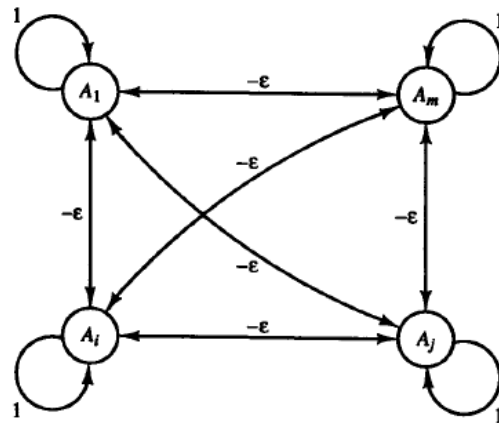


**Gambar 2.4.** Jaringan Layer Jamak

Gambar 2.4 adalah jaringan dengan  $i$  *neuron input* ( $x_1, x_2, \dots, x_n$ ), sebuah layer tersembunyi yang terdiri dari  $j$  *neuron* ( $\varphi_1, \varphi_2, \dots, \varphi_k$ ) dan  $s$  *neuron output* ( $Y_1, Y_2, \dots, Y_m$ ). Jaringan ini dapat menyelesaikan masalah yang lebih kompleks dibandingkan dengan layer tunggal, meskipun kadangkala proses pelatihan lebih kompleks dan lama.

c. Jaringan Layer Kompetitif (*competitive layer network*)

Arsitektur ini memiliki bentuk yang berbeda, dimana antar *neuron* dapat saling dihubungkan. Gambar 2.5 merupakan salah satu contoh arsitektur ini.



**Gambar 2.5.** Jaringan Layer Kompetitif

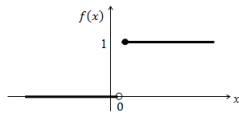
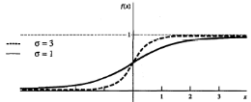
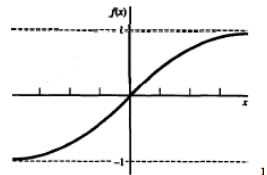
## 2. Fungsi Aktivasi

Dalam *Neural Network*, fungsi aktivasi digunakan untuk menentukan *output* suatu *neuron*. Berikut merupakan tabel 4 fungsi aktivasi dalam *Neural Network* menurut Fausett (1994: 17-19):

**Tabel 2.1** Fungsi Aktivasi *Neural Network*

No	Nama	Pengertian	Fungsi	Grafik
1.	Fungsi Identitas ( <i>Identity Function</i> )	Pada fungsi identitas, nilai <i>output</i> yang dihasilkan sama dengan nilai <i>inputnya</i>	$y = f(x) = x,$ $x \in R$	



2.	Undak Biner ( <i>Binary Step</i> )	Jaringan dengan lapisan tunggal sering menggunakan fungsi undak ( <i>step function</i> ) untuk mengkonversikan <i>input</i> dari suatu variabel yang bernilai kontinu ke suatu <i>output</i> biner (0 atau 1)	$y = \begin{cases} 0, & \text{jika } x < 0 \\ 1, & \text{jika } x \geq 0 \end{cases}$	
3.	Fungsi Sigmoid Biner	Fungsi ini digunakan untuk <i>Neural Network</i> yang dilatih dengan menggunakan metode <i>backpropagation</i> . Fungsi sigmoid biner memiliki nilai pada range 0 sampai 1, sehingga sering digunakan untuk <i>Neural Network</i> yang membutuhkan nilai <i>output</i> yang terletak pada interval 0 sampai 1. Namun, fungsi ini bisa juga digunakan oleh <i>Neural Network</i> yang nilai <i>output</i> nya 0 atau 1	$y = f(x) = \frac{1}{1 + e^{-\sigma x}}$ <p>dengan:</p> $f'(x) = \sigma f(x)[1 - f(x)]$	
4.	Fungsi Sigmoid Bipolar ( <i>Bipolar Sigmoid</i> )	Fungsi sigmoid bipolar berkaitan dengan fungsi tangen hiperbolik yang sering digunakan sebagai fungsi aktivasi ketika nilai <i>output</i> yang dibutuhkan terletak pada interval -1 sampai 1	$y = f(x) = \frac{1 - e^{-\sigma x}}{1 + e^{-\sigma x}}$ <p>dengan:</p> $f'(x) = \frac{\sigma}{2} [1 + f(x)][1 - f(x)]$	

### 3. Algoritma Pembelajaran (*learning algorithm*)

Algoritma pembelajaran adalah prosedur untuk menentukan bobot pada lapisan yang berhubungan dalam *Neural Network* (Fausett, 1994: 429). Selama proses pembelajaran akan terjadi perbaikan bobot-bobot berdasarkan algoritma tertentu. Nilai bobot akan bertambah, jika informasi yang diberikan oleh *neuron* yang bersangkutan tersampaikan, sebaliknya jika informasi tidak tersampaikan oleh suatu *neuron* ke *neuron* lain, maka nilai bobot yang menghubungkan keduanya akan dikurangi. Pada saat pembelajaran dilakukan pada *input* yang berbeda maka nilai bobot akan diubah secara dinamis hingga mencapai suatu nilai yang cukup seimbang. Apabila nilai ini telah seimbang, maka mengindikasikan bahwa tiap-tiap input telah berhubungan dengan *output* yang diharapkan (Sri Kusumadewi & Sri Hartati, 2010).

Terdapat dua metode yang dapat dilakukan dalam proses pembelajaran NN, yaitu metode pembelajaran terawasi (*supervised learning*) dan metode tak terawasi (*unsupervised learning*) (Edy Irwansyah & M. Faisal, 2015: 53):

#### a. *Supervised Learning*

Metode pembelajaran pada *Neural Network* disebut terawasi jika *output* yang diharapkan telah diketahui terlebih dahulu. Pengetahuan yang akan diberikan kepada sistem awalnya diberikan suatu acuan untuk memetakan suatu *input* menjadi *output* yang diinginkan. Proses pembelajaran ini akan terus dilakukan selama kondisi *error* atau kondisi yang diinginkan belum tercapai. Adapun setiap perolehan *error* akan

dikalkulasikan untuk setiap pemrosesan hingga data atau nilai yang diinginkan telah tercapai.

b. *Unsupervised Learning*

Sistem pembelajaran tidak terawasi memerlukan suatu target *output*. Pada sistem ini tidak membutuhkan adanya acuan awal, agar perolehan nilai dapat dicapai. Pada metode ini hasil yang diharapkan selama proses pembelajaran tidak dapat ditentukan. Tujuan pembelajaran *unsupervised* adalah mengelompokkan unit-unit yang hampir sama dalam area tertentu.

**E. Metode *Clustering***

Dalam proses pengelompokan data (*clustering*), sebelumnya ditentukan nilai suatu jarak untuk mengukur kemiripan dari objek-objek yang diamati. Jarak yang umumnya digunakan yaitu jarak *Euclidean*. Semakin kecil nilai jarak *Euclidean*, semakin tinggi tingkat kemiripan, begitu pula sebaliknya, semakin besar nilai jarak *Euclidean* maka semakin rendah tingkat kemiripannya. Setelah ukuran kemiripan ditemukan, maka dapat dilakukan pengelompokan (Brodjolo Sutijo, 2008). Terdapat beberapa metode yang dapat digunakan, diantaranya adalah sebagai berikut.

**1. K-Means *Clustering***

Algoritma K-Means *clustering* dikembangkan oleh MacQueen (1967) kemudian Hartigan dan Wong sekitar tahun 1975. Sederhananya, K-Means merupakan algoritma untuk mengklasifikasikan atau mengelompokkan objek/data berdasarkan unsur/fitur ke sejumlah  $k$  kelompok/*cluster*, dengan  $k$

adalah bilangan bulat positif. Pengelompokan dilakukan dengan meminimalkan jumlah kuadrat dari jarak data dengan pusat *cluster* yang sesuai (Teknomo, 2015).

Algoritma metode K-Means *clustering* adalah sebagai berikut (Johnson & Wichern, 2007: 696):

- a. Partisi data ke dalam  $k$  *cluster*
- b. Tempatkan setiap data/obyek ke kluster terdekat. Kedekatan dua obyek ditentukan berdasarkan jarak kedua obyek tersebut. Jarak biasanya dihitung dengan menggunakan jarak *Euclide*. Persamaan jarak *Euclide* antara dua titik sebarang P dan Q dengan koordinat P  $(x_1, x_2, \dots, x_n)$  dan Q  $(y_1, y_2, \dots, y_n)$  adalah sebagai berikut:

$$d(P, Q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (2.16)$$

Hitung ulang nilai pusat untuk *cluster* yang menerima data baru dan *cluster* yang kehilangan data

- c. Ulangi langkah ke-2 sampai nilai pusat lama sama dengan nilai pusat baru (stabil).

Beberapa keunggulan K-Means *clustering* antara lain (Zhang C & Fang Z, 2013):

- a. Algoritma K-Means merupakan algoritma klasik untuk menyelesaikan masalah pengelompokkan. Algoritma ini relatif sederhana dan cepat.
- b. Untuk data yang besar, algoritma ini relatif fleksibel dan efisien.
- c. Memberikan hasil yang relatif baik

Beberapa kekurangan K-Means *clustering* antara lain (Zhang C & Fang Z, 2013):

- a. Sensitif terhadap nilai awal, sehingga apabila nilai awal berbeda, mungkin akan terbentuk *cluster* yang berbeda.
- b. Algoritma K-Means *clustering* memiliki ketergantungan yang lebih tinggi dari pusat *cluster* awal. Jika pusat *cluster* awal benar-benar jauh dari pusat *cluster* data itu sendiri, jumlah iterasi cenderung tak terbatas dan menghasilkan pengelompokan yang tidak tepat.
- c. Algoritma K-Means *clustering* memiliki sensitifitas yang kuat terhadap *noise* objek data. Jika terdapat sejumlah data *noise* pada kumpulan data, ini akan mempengaruhi hasil pengelompokan akhir yang menyebabkan *error* pada hasil.

## 2. Fuzzy C-Means (FCM) *Clustering*

Fuzzy C-Means (FCM) merupakan salah satu algoritma fuzzy *clustering*. Metode ini merupakan pengembangan dari metode *non hierarkhi K-Means Cluster*, karena pada awalnya ditentukan dulu jumlah kelompok atau *cluster* yang akan dibentuk berbeda dengan metode *hierarkhi* yang jumlah kelompoknya akan membentuk semacam ‘pohon’ dimana terdapat tingkatan yang jelas antara objek, dari yang paling mirip hingga yang paling tidak mirip (Soleh dkk, 2014). Fuzzy C-Means merupakan suatu teknik pengelompokan data yang keberadaan setiap titik data dalam suatu *cluster* ditentukan oleh derajat keanggotaan. Fungsi objektif yang digunakan adalah (Zimmerman, 2001:294) dan (Yang & Huang, 2007):

$$J_{FCM} = \sum_{j=1}^k \sum_{s=1}^m (\mu_{sj})^w d_{sj} \quad (2.17)$$

Dimana:

$\mu_{sj}$  = derajat keanggotaan data ke- $s$  ( $x_s$ ) pada *cluster* ke- $j$ , yang dapat dicari dengan Persamaan

$$\mu_{sj} = \frac{1}{\sum_{l=1}^k \left(\frac{d_{sj}}{d_{sl}}\right)^{\frac{1}{w-1}}} \quad (2.18)$$

$k$  = banyak *cluster* yang memenuhi (jumlah *cluster* yang diinginkan),

$$2 \leq k < q$$

$w$  = pangkat atau bobot,  $1 < w < \infty$

$m$  = jumlah data (observasi)

$d_{sj}$  = jarak *Euclidean* antara data ke- $s$  ( $x_s$ ) dengan pusat *cluster* ke- $j$

$d_{sl}$  = jarak *Euclidean* antara data ke- $s$  ( $x_s$ ) dengan pusat *cluster* ke- $l$

Algoritma pengelompokkan dengan menggunakan Fuzzy C-Means adalah sebagai berikut:

- a. Input data yang akan dikelompokkan, berupa matriks berukuran  $s \times i$   
( $s$  = jumlah data,  $i$  = variabel data)
- b. Tentukan jumlah *cluster* ( $k$ ), pangkat atau bobot ( $w$ ), *error* terkecil yang diharapkan ( $\xi$ )
- c. Menetapkan matriks partisi awal

$$U_{(0)} = \begin{bmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{1j} \\ \mu_{21} & \mu_{22} & \dots & \mu_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{s1} & \mu_{s2} & \dots & \mu_{sj} \end{bmatrix} \quad (2.19)$$

- d. Menghitung pusat *cluster* ke- $j$ :  $V_{ji}$ , dengan  $j = 1, 2, \dots, k$ ; dan  $i = 1, 2, \dots, n$  menggunakan Persamaan berikut:

$$V_{ji} = \frac{\sum_{s=1}^m (\mu_{sj})^w x_{si}}{\sum_{s=1}^m (\mu_{si})^w} \quad (2.20)$$

Dengan,

$V_{ji}$  = pusat *cluster* ke-  $j$  untuk variabel ke-  $i$ ,

$\mu_{sj}$  = derajat keanggotaan data ke-  $s$  pada *cluster* ke-  $j$ ,

$w$  = pangkat atau bobot,

$x_{si}$  = data ke-  $s$  pada variabel ke-  $i$ ,

Dalam mencari pusat *cluster*, dapat menggunakan MATLAB R2013a dengan fungsi (Elena, 2013):

$$[\text{center}, U, \text{obj\_fcn}] = \text{fcm}(\text{data}, \text{n\_clusters}) \quad (2.21)$$

dengan,

`center` : matriks pusat *cluster*

`U` : matriks derajat keanggotaan

`obj_fcn` : nilai fungsi *objective* pada setiap iterasi

`data` : data yang akan dikelompokkan

`n_clusters` : banyak *cluster* yang akan dibentuk (lebih dari satu).

- e. Menentukan nilai fungsi objektif pada iterasi ke-  $t$ ; yaitu  $J_{(t)}$ , dengan menggunakan Persamaan (2.21)

- f. Menentukan kriteria berhenti, yaitu perubahan matriks partisi pada iterasi sekarang dengan iterasi sebelumnya, jika

$$|J_{(t+1)} - J_{(t)}| < \xi \quad (2.22)$$

Maka berhenti. Jika tidak, iterasi dinaikkan ( $t = t + 1$ ) dan ulangi langkah ke-4 (d).

- g. Setelah mendapatkan pusat *cluster* masing-masing data, selanjutnya mencari jarak *Euclide* dengan rumus (Indira & Can, 2013):

$$d_{js} = d(X_s - V_j) = \sqrt{\sum_{i=1}^n (X_{si} - V_{ji})^2} \quad (2.23)$$

dengan,

$d_{js}$  = jarak *Euclide cluster* ke- $j$  data ke- $s$

$X_{si}$  = data ke- $s$  variabel ke- $i$

$V_{ji}$  = pusat *cluster* ke- $j$  variabel ke- $i$

$i = 1, 2, \dots, n$

$s = 1, 2, \dots, m$

$j = 1, 2, \dots, k$

Setelah mendapatkan jarak *Euclide* dari masing-masing data, selanjutnya mencari jarak maksimum dari masing-masing *cluster* nya sebelum dilakukan penghitungan bobot lapisan output jaringan optimum.

Beberapa keunggulan Fuzzy C-Means *clustering* antara lain (Dang Q. A. *et al*, 2015):

- a. Memberikan hasil yang terbaik untuk kumpulan data yang *overlap*.



- b. Ditetapkan derajat keanggotaan untuk setiap pusat *cluster*, akibatnya memungkinkan untuk titik data masuk ke lebih dari satu pusat *cluster*.

Beberapa kekurangan Fuzzy C-Means *clustering* antara lain (Dang Q. A. *et al*, 2015):

- a. Lebih banyak proses iterasi dan perhitungan yang lama
- b. Sensitif terhadap inisialisasi jumlah *cluster*

## **F. Ridge Regression**

Sekitar pertengahan abad ke-20 teoritikus Rusia Andre Tikhonov mengerjakan solusi dari masalah *ill-posed*. Ini adalah kasus matematika yang tidak mempunyai solusi, karena pada dasarnya tidak ada cukup informasi khusus dalam kasus tersebut. Hal ini diperlukan untuk memberikan informasi tambahan (atau asumsi) sehingga teknik matematika Tikhonov dikembangkan untuk kasus ini yang dikenal sebagai regularisasi.

Kerja Tikhonov menjadi dikenal secara luas di Barat setelah publikasi bukunya pada tahun 1997. Sementara itu, dua ahli statistik Amerika, Arthur Hoerl dan Robert Kennard, menerbitkan sebuah makalah pada tahun 1970 pada *ridge regression*, metode untuk memecahkan masalah regresi linear yang buruk. Kondisi buruk berarti kesulitan numerik dalam menjalankan matriks inverse yg diperlukan untuk mendapatkan matriks variansi. Hal ini merupakan gejala dari masalah regresi *ill-posed* dalam pengertian Tikhonov dan metode Hoerl & Kennard adalah bentuk dasar dari regularisasi, yang sekarang dikenal sebagai regularisasi orde nol.

Pada tahun 1980-an, ketika jaringan saraf menjadi populer, *weight decay* adalah salah satu dari sejumlah teknik 'temuan' untuk membantu pangkasan koneksi jaringan yang tidak penting. Namun, segera diakui bahwa *weight decay* melibatkan penambahan *penalty* yang sama dengan *Sum Squared Error* seperti dalam *ridge regression*. *Weight decay* ekuivalen dengan *ridge regression*. Sementara itu *ridge regression* secara matematis dan komputasi mudah digunakan dan akibatnya bentuk lain dari regularisasi cukup diabaikan.

Selanjutnya dijelaskan mengenai *ridge regression* dari perspektif bias dan variansi sebagaimana hal itu mempengaruhi persamaan untuk vektor bobot yang optimal, matriks variansi, dan matriks proyeksi.

### 1. Bias dan Variansi

Ketika input  $x$  model train memprediksi output sebagai  $f(x)$ . Jika terdapat banyak kumpulan data training (yang tidak pernah dilakukan, hanya perkiraan) dan jika diketahui output  $y(x)$  yang benar. Dapat dihitung *Mean Squared Error* (MSE), yaitu

$$MSE = \langle (y(x) - f(x))^2 \rangle \quad (2.24)$$

Nilai ini, menunjukkan seberapa baik prediksi rata-rata, yang dapat dipecah menjadi dua komponen yaitu

$$MSE = (y(x) - \langle f(x) \rangle)^2 + \langle (f(x) - \langle f(x) \rangle)^2 \rangle \quad (2.25)$$

Bagian pertama adalah bias dan bagian kedua adalah variansi.

Jika  $\langle f(x) \rangle = y(x)$  untuk semua  $x$  maka model ini tidak bias (bias bernilai nol). Namun, model tidak bias mungkin masih memiliki *Mean Squared Error* bernilai besar jika memiliki variansi yang bernilai besar. Ini akan terjadi

jika  $f(x)$  sangat sensitif terhadap kekhasan (seperti *noise* dan pilihan titik sampel) dari setiap kumpulan data training tertentu, sensitivitas ini yang menyebabkan masalah regresi menjadi *ill-posed* dalam pengertian Tikhonov. Variansi dapat dikurangi secara signifikan dengan memasukkan sejumlah kecil bias sehingga terjadi pengurangan *Mean Squared Error*.

Masuknya bias setara dengan pembatasan jangkauan pada fungsi dimana model dapat dijelaskan. Hal ini dicapai dengan menghapus derajat kebebasan. Misalnya akan menurunkan urutan polinomial atau mengurangi jumlah bobot dalam *Neural Network*, *ridge regression* menghapus derajat kebebasan tidak secara eksplisit tetapi mengurangi jumlah efektif dari parameter. Hasilnya berupa hilangnya fleksibilitas yang membuat model kurang sensitif.

$$C = \sum_{i=1}^p (\hat{y}_i - f(x))^2 + \lambda \sum_{j=1}^m w_j^2 \quad (2.26)$$

Ini merupakan *ridge regression (weight decay)* dengan  $\hat{y}_i$  merupakan nilai variabel output ke- $i$  dengan  $p$  adalah banyaknya pengamatan dan  $w_j$  adalah bobot dari *neuron* lapisan tersembunyi ke- $j$ . Parameter regularisasi  $\lambda > 0$  mengatur keseimbangan antara penyesuaian data dan pencegahan *penalty*.  $\lambda$  bernilai kecil menunjukkan bahwa data tersebut tepat tanpa menyebabkan *penalty* bernilai besar. Sedangkan  $\lambda$  bernilai besar menunjukkan ketepatan data tidak bisa didapatkan jika membutuhkan bobot besar. Bias merupakan solusi yang melibatkan bobot bernilai kecil dan hasilnya untuk proses fungsi *output* karena bobot yang besar diperlukan untuk menghasilkan fungsi *output* yang sangat bervariasi (*rough*). (Orr, 1996: 23-24).

Berdasarkan pernyataan-pernyataan sebelumnya, diketahui bahwa masalah yang mungkin muncul ketika bekerja dengan *noise* pada data training, input dalam jumlah besar, dan kumpulan training dalam jumlah kecil. Hal ini disebut *over-fitting*. Dalam mengatasi masalah tersebut, sebuah *roughness penalty*, yaitu ukuran kemulusan kurva dalam memetakan data, dapat ditambahkan pada *Sum Square Error* (SSE). Ini yang disebut *global ridge regression*. Metode *global ridge regression* mengestimasi bobot dengan menambahkan parameter regulasi tunggal yang bernilai positif pada *Sum Square Error* (SSE) untuk mendapatkan vektor bobot yang lebih kuat terhadap *noise* pada kumpulan data training (Leondes C.T, 2005: 128).

Kriteria pemilihan model mencakup estimasi prediksi *error*, yaitu estimasi seberapa baik model pada data *training* akan bekerja pada input selanjutnya yang tidak diketahui. Model yang terbaik adalah model dengan estimasi prediksi *error* yang kecil. Salah satu kriteria tersebut yaitu *Generalised Cross-Validation* (GCV) untuk menghitung prediksi *error*. Rumus GCV adalah sebagai berikut. (Orr, 1996: 20).

$$\hat{\sigma}_{GCV}^2 = \frac{m\hat{y}^T P^2 \hat{y}}{(\text{trace}(P))^2} \quad (2.27)$$

$m$  = banyak data

$$P = I_m - \varphi A^{-1} \varphi^T$$

$P^2$  = matriks proyeksi

$\hat{y}$  = vektor target klasifikasi

## 2. Pengoptimalan Parameter Regulasi

Pemilihan model digunakan untuk memilih nilai untuk parameter regularisasi  $\lambda$ . Nilai yang dipilih adalah salah satu yang terkait dengan estimasi *error* terendah.

Karena semua kriteria pemilihan model bergantung secara non linear pada  $\lambda$  dibutuhkan metode optimasi nonlinier. Sehingga digunakan salah satu teknik standar untuk ini, seperti ketika turunan dari estimasi *error* GCV disamadengankan nol, persamaan yang dihasilkan dapat dimanipulasi sehingga hanya  $\lambda$  yang muncul di sisi kiri.

$$\hat{\lambda} = \frac{\hat{y}^T P^2 \hat{y} \text{trace} (A^{-1} - \hat{\lambda} A^{-2})}{\hat{w}^T A^{-1} \hat{w} \text{trace} (P)} \quad (2.28)$$

Persamaan tersebut bukan solusi, namun merupakan rumus estimasi ulang karena sisi kanan bergantung pada  $\hat{\lambda}$  (secara eksplisit maupun implisit melalui  $A^{-1}$  dan  $P$ ). Untuk menggunakannya, nilai awal dari  $\hat{\lambda}$  dipilih dan digunakan untuk menghitung nilai untuk sisi kanan, ini menyebabkan perkiraan baru dan proses dapat diulang sampai konvergen.

### G. Ketepatan Hasil Klasifikasi

Setelah proses pembelajaran selesai dilakukan, tahapan selanjutnya adalah pengujian ketepatan hasil klasifikasi (diagnosa). Sensitivitas, spesifisitas dan akurasi secara luas digunakan untuk menggambarkan hasil klasifikasi. Secara khusus, digunakan untuk mengukur seberapa baik dan terpercaya hasil klasifikasi tersebut. Kemungkinan yang dapat terjadi pada hasil klasifikasi ditunjukkan Tabel 2.2 berikut (Zhu Wen *et al*, 2010).

**Tabel 2.2.** Hasil Klasifikasi Uji Diagnosa

Hasil Diagnosa	Kondisi penyakit sebagaimana ditetapkan oleh Standar Kebenaran		
	Positif	Negatif	Jumlah Baris
Positif	TP	FP	TP+FP
Negatif	FN	TN	FN+TN
Jumlah Kolom	TP+FN	FP+TN	N = TP+TN+FP+FN

Ada beberapa istilah yang umum digunakan bersama dengan deskripsi sensitivitas, spesifisitas dan akurasi yaitu TP = *True Positive*, FP = *False Positive*, TN = *True Negative*, dan FN = *False Negative*

Jika penyakit terbukti ada dalam tubuh pasien, tes diagnostik yang diberikan juga menunjukkan adanya penyakit, hasil tes diagnostik dianggap *True Positive* (TP). Demikian pula, jika penyakit terbukti tidak ada pada tubuh pasien, tes diagnostik menunjukkan penyakit tidak ada juga, sehingga hasil tes *True Negative* (TN). Kedua *True Negative* and *True Positive* menunjukkan hasil yang konsisten antara tes diagnostik dan kondisi terbukti (juga disebut standar kebenaran). Namun, tidak ada tes medis yang sempurna.

Jika tes diagnostik menunjukkan adanya penyakit pada pasien yang sebenarnya tidak memiliki penyakit tersebut, hasil tes *False Positive* (FP). Demikian pula, jika hasil tes diagnosis menunjukkan bahwa penyakit ini tidak ada pada pasien yang sebenarnya terjangkit penyakit, hasil tes *False Negative* (FN). Kedua *False Positive* dan *False Negative* menunjukkan bahwa hasil tes berlawanan dengan kondisi yang sebenarnya. (Zhu Wen *et al*, 2010).

## 1. Sensitivitas

Sensitivitas mengacu pada kemampuan tes untuk mengidentifikasi pasien dengan penyakit secara tepat. Rumus Sensitivitas adalah sebagai berikut.

$$\text{sensitivitas} = \frac{TP}{TP+FN} \quad (2.29)$$

Misalnya, jika sensitivitas = 99%, artinya ketika dilakukan tes diagnostik pada pasien dengan penyakit tertentu, pasien ini berpeluang 99% teridentifikasi positif terjangkit penyakit tersebut.

## 2. Spesifisitas

Spesifisitas mengacu pada kemampuan tes untuk mengidentifikasi pasien tanpa penyakit secara tepat. Rumus Spesifisitas adalah sebagai berikut.

$$\text{spesifisitas} = \frac{TN}{TN+FP} \quad (2.30)$$

Nilai Spesifisitas merupakan peluang tes diagnosa penyakit tertentu tanpa memberikan hasil *False Positive*. Misalnya, jika spesifisitas suatu tes 99%, ini artinya ketika dilakukan tes diagnosa pada pasien tanpa penyakit tertentu, pasien ini berpeluang 99% teridentifikasi negatif terjangkit penyakit tersebut

## 3. Akurasi

Akurasi adalah proporsi dari hasil yang benar (True), baik *True Positive* maupun *True Negative*, dalam suatu populasi. Akurasi mengukur atau mengidentifikasi dengan benar kondisi pasien. Rumus untuk menghitung akurasi adalah sebagai berikut.

$$\text{Akurasi} = \frac{TN+TP}{TN+TP+FN+FP} = \frac{\text{Jumlah diagnosis benar (cocok)}}{\text{Jumlah diagnosis keseluruhan}} \quad (2.31)$$

Besar nilai akurasi merepresentasikan tingginya keakuratan hasil diagnosa pada pasien yang melakukan uji diagnosa, baik pasien yang terjangkit penyakit maupun tidak.