

NAMA : RADEN ISNAWAN ARGY ARYASATYA  
NIM : 195410257  
KELAS : IF-5  
MATKUL : DATA MINING

---

### LATIHAN SOAL (PERTEMUAN 13)

1. Kerjakan soal yang ada di materi Pertemuan 8 (slide terakhir).

Diketahui:

Jika diketahui dari data kelulusan mahasiswa disebut perguruan tinggi adalah sebagai berikut:

NO	JENIS KELAMIN	STATUS MAHASISWA	STATUS PRENIKAHAN	IPK Semester 1-6	STATUS KELULUSAN
1	LAKI - LAKI	MAHASISWA	BELUM	3.17	TEPAT
2	LAKI - LAKI	BEKERJA	BELUM	3.30	TEPAT
3	PEREMPUAN	MAHASISWA	BELUM	3.01	TEPAT
4	PEREMPUAN	MAHASISWA	MENIKAH	3.25	TEPAT
5	LAKI - LAKI	BEKERJA	MENIKAH	3.20	TEPAT
6	LAKI - LAKI	BEKERJA	MENIKAH	2.50	TERLAMBAT
7	PEREMPUAN	BEKERJA	MENIKAH	3.00	TERLAMBAT
8	PEREMPUAN	BEKERJA	BELUM	2.70	TERLAMBAT
9	LAKI - LAKI	BEKERJA	BELUM	2.40	TERLAMBAT
10	PEREMPUAN	MAHASISWA	MENIKAH	2.50	TERLAMBAT
11	PEREMPUAN	MAHASISWA	BELUM	2.50	TERLAMBAT
12	PEREMPUAN	MAHASISWA	BELUM	3.50	TEPAT
13	LAKI - LAKI	BEKERJA	MENIKAH	3.30	TEPAT
14	LAKI - LAKI	MAHASISWA	MENIKAH	3.25	TEPAT
15	LAKI - LAKI	MAHASISWA	BELUM	2.30	TERLAMBAT

## TUGAS

Jika seorang mahasiswa dengan data sebagai berikut, buatlah prediksi kelulusannya!

KELAMIN	STATUS	PRENIKAHAN	IPK	KETERANGAN
PEREMPUAN	BEKERJA	BELUM	3.15	???

Jawab:

- Tahap 1: menghitung jumlah class/label
  - $P(Y=TEPAT) = 8/15$   
(Jumlah data "TEPAT" pada kolom "STATUS KELULUSAN" dibagi jumlah data)

- $P(Y=\text{TERLAMBAT}) = 7/15$   
(Jumlah data "TERLAMBAT" pada kolom "STATUS KELULUSAN" dibagi dengan semua jumlah data)
- Tahap 2: menghitung jumlah kasus yang sama dengan class yang sama
  - $P(\text{JENIS KELAMIN} = \text{PEREMPUAN} \mid Y = \text{TEPAT}) = 3/8$   
(jumlah data jenis kelamin "PEREMPUAN" dengan keterangan "TEPAT" dibagi jumlah data TEPAT)
  - $P(\text{JENIS KELAMIN} = \text{PEREMPUAN} \mid Y = \text{TERLAMBAT}) = 4/7$   
(jumlah data jenis kelamin "PEREMPUAN" dengan keterangan "TERLAMBAT" dibagi jumlah data TERLAMBAT)
  - $P(\text{STATUS MAHASISWA} = \text{BEKERJA} \mid Y = \text{TEPAT}) = 3/8$   
(jumlah data dengan status "BEKERJA" dengan keterangan "TEPAT" dibagi jumlah data "TEPAT")
  - $P(\text{STATUS MAHASISWA} = \text{BEKERJA} \mid Y = \text{TERLAMBAT}) = 4/7$   
(jumlah data dengan status "BEKERJA" dengan keterangan "TERLAMBAT" dibagi jumlah data "TERLAMBAT")
  - $P(\text{STATUS PERNIKAHAN} = \text{BELUM} \mid Y = \text{TEPAT}) = 4/8$   
(jumlah data dengan status pernikahan "BELUM" dan keterangan "TEPAT" dibagi jumlah data TEPAT)
  - $P(\text{STATUS PERNIKAHAN} = \text{BELUM} \mid Y = \text{TERLAMBAT}) = 4/7$   
(jumlah data dengan status pernikahan "BELUM" dan keterangan "TERLAMBAT" dibagi jumlah data TERLAMBAT)
  - $P(\text{IPK} = 3.15 \mid Y = \text{TEPAT}) = 0/8$   
(jumlah data IPK "3.15" dengan keterangan "TEPAT" dibagi jumlah data TEPAT)
  - $P(\text{IPK} = 3.15 \mid Y = \text{TERLAMBAT}) = 0/7$   
(jumlah data IPK "3.15" dengan keterangan "TERLAMBAT" dibagi jumlah data TERLAMBAT)
- Tahap 3: kalikan semua hasil variabel TEPAT dan variabel TERLAMBAT
  - TEPAT:  

$$(P(\text{KELAMIN} = \text{PEREMPUAN} \mid Y = \text{TEPAT}) * P(\text{STATUS MAHASISWA} = \text{BEKERJA} \mid Y = \text{TEPAT}) * P(\text{PERNIKAHAN} = \text{BELUM} \mid Y = \text{TEPAT}) * P(\text{IPK} = 3.15 \mid Y = \text{TEPAT}))$$

$$= 3/8 * 3/8 * 4/8 * 0/8$$

$$= 0$$
  - TERLAMBAT:  

$$(P(\text{KELAMIN} = \text{PEREMPUAN} \mid Y = \text{TERLAMBAT}) * P(\text{STATUS MAHASISWA} = \text{BEKERJA} \mid Y = \text{TERLAMBAT}) * P(\text{PERNIKAHAN} = \text{BELUM} \mid Y = \text{TERLAMBAT}) * P(\text{IPK} = 3.15 \mid Y = \text{TERLAMBAT}))$$

$$= 4/7 * 4/7 * 4/7 * 0/7$$

$$= 0$$
- Tahap 4: kalikan status kelulusan TEPAT dan TERLAMBAT dengan jumlah TEPAT dan TERLAMBAT
  - TEPAT:  

$$(P(\text{STATUS KELULUSAN} = \text{"TEPAT"}) * (\text{JUMLAH TEPAT})) = 0 * 8/15 = 0$$

➤ TERLAMBAT:

➤  $(P(\text{STATUS KELULUSAN} = \text{"TERLAMBAT"}) * (\text{JUMLAH TERLAMBAT})) = 0 * 7/15 = 0$

- Tahap 5: Tambahkan 1 data pura-pura untuk data "TEPAT" dan "TERLAMBAT".  
Karena hasil akhir status kelulusan sama-sama 0, maka kita tambahkan satu data dummy (data pura-pura) dengan IPK 3.15 dengan menambahkan nilai 1 pada tiap variabel untuk menentukan hasil akhir.

➤ TEPAT:

$$4/8 * 4/8 * 5/8 * 1/8 = 0,0195312$$

➤ TERLAMBAT

$$5/7 * 5/7 * 5/7 * 1/7 = 0,0520616$$

- Tahap 6: hasil akhir / kesimpulan

Data *JENIS KELAMIN* = PEREMPUAN / *STATUS MAHASISWA* = BEKERJA / *PERNIKAHAN* = BELUM / *IPK* = 3.15 masuk ke kelas "TERLAMBAT"

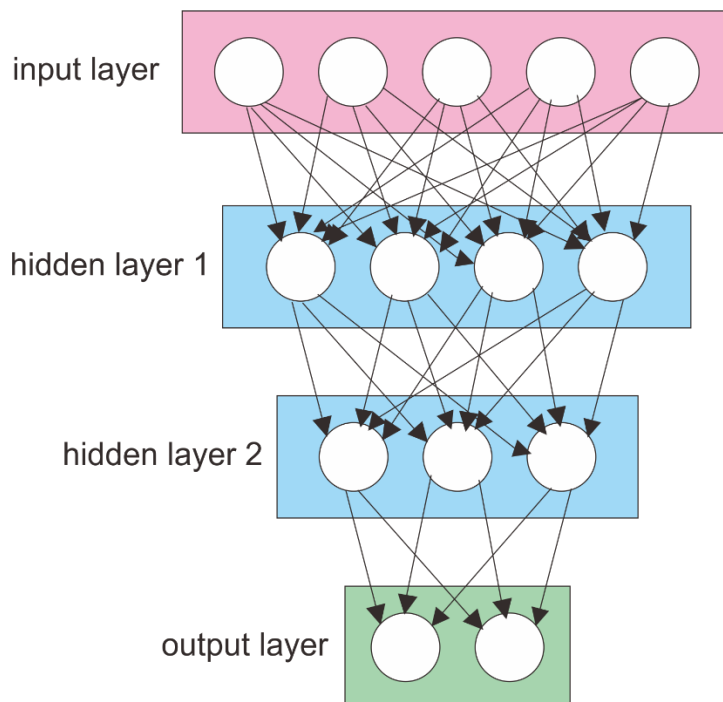
2. Gambarkan Jaringan saraf tiruan multilayer yang terdiri dari 5 unit input, 2 hidden layer yang masing-masing terdiri dari 4 neuron dan 3 neuron, serta 2 unit output.

Jawab:

Gambar dibawah saya gambar menggunakan CorelDraw X7.

Klik link berikut jika ingin lihat atau download file CorelDraw dan PNG nya:

[https://drive.google.com/drive/folders/1jsrcyqgN-NGS5HbUedjDqG8\\_WZvKg2zR?usp=sharing](https://drive.google.com/drive/folders/1jsrcyqgN-NGS5HbUedjDqG8_WZvKg2zR?usp=sharing)



3. Suatu neuron j menerima masukan dari 5 neuron lain dengan input 4, 3, -1, 5 dan 2. Bobot dari kelima neuron tersebut ke j masing-masing 0,2; 0,4; 0,3;-0,1 dan -0,3. Hitunglah output neuron j jika fungsi aktivasi yang digunakan pada neuron j adalah : fungsi aktivasi bipolar dengan nilai threshold 2

Jawab:

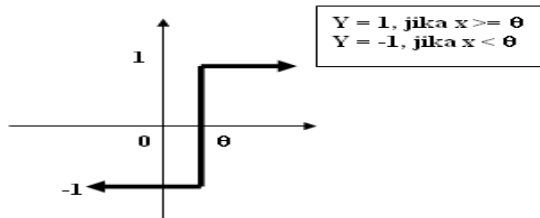
Fungsi bipolar sebenarnya hampir sama dengan fungsi undak biner dengan threshold, hanya saja output yang dihasilkan berupa 1, 0 atau -1

Fungsi bipolar (dengan nilai ambang 0) dirumuskan sebagai:

$Y = 1$ , jika  $x \geq 0$

$Y = -1$ , jika  $x < 0$

Rumus:



Input: 4, 3, -1, 5, 2

Bobot: 0,2; 0,4; 0,3; -0,1 dan -0,3

Berarti, ada 5 variabel input yaitu  $x_1, x_2, x_3, x_4, x_5$ . Saat disambungkan ke suatu node, tiap node input wajib memiliki bobot (weight) yang awalnya ditentukan secara random (sekarang sudah diberi nilai yaitu 0,2; 0,4; 0,3; -0,1 dan -0,3) saat masuk ke dalam node semua nilai  $x$  dan bobot harus dikalikan kemudian diagregasi.

Nilai input node :  $(x_1 * w_1 + b_1) + (x_2 * w_2 + b_2) + (x_3 * w_3 + b_3) + (x_4 * w_4 + b_4) + (x_5 * w_5 + b_5)$

$b$  adalah nilai bias. Nilai bias boleh digunakan boleh tidak. Nilai bias biasanya digunakan apabila nilai bobot atau nilai input 0 sehingga dikalikan menjadi nol dan tidak teraktivasi. Nilai  $b$  biasanya 1 untuk menghindari node tidak aktif.

Jadi, perhitungan nilai input node J adalah seperti berikut:

$(x_1 * w_1) + (x_2 * w_2) + (x_3 * w_3) + (x_4 * w_4) + (x_5 * w_5)$

Nilai node J  
 $= (4 * 0,2) + (3 * 0,4) + (-1 * 0,3) + (5 * -0,1) + (2 * -0,3)$   
 $= 0,8 + 1,2 + (-0,3) + (-0,5) + (-0,6)$   
 $= 0,6$

Nilai node J diatas masih belum diaktivasi, fungsi aktivasi digunakan untuk mengubah nilai hasil hitung bobot menjadi nilai range 0–1 atau -1–1 karena semua nilai yang disimpan di node harus dalam range tersebut

Nilai output dengan aktivasi:

Nilai output J  $= f(x) = f(0,6) = 0,9$   
 $= (0,9 * 0,2) + (0,9 * 0,4) + (0,9 * 0,3) + (0,9 * -0,1) + (0,9 * -0,3)$   
 $= 0,18 + 0,36 + 0,27 + (-0,09) + (-0,27)$   
 $= 0,45 = f(x) = f(0,45)$

Nilai output J  $= 0,675$

Nilai error output J  $= 1 - 0,675 = 0,325$

Jadi, output nilai J adalah 0,675 dengan error sebesar 0,325.