



Data Mining

Fungsi Estimasi

Tutik Khotimah, S.Kom, M.Kom

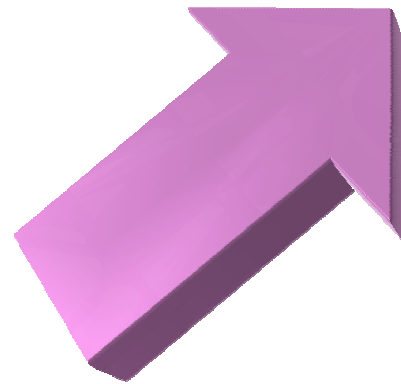


Tujuan Instruksional

Menjelaskan Fungsi Estimasi

**Menjelaskan Regresi Linear
Sederhana**

**Menjelaskan Regresi Linear
Berganda**



**Menjelaskan
Fungsi Estimasi**

Fungsi Estimasi





Fungsi Estimasi

- ☐ Termasuk dalam metode learning Supervised Learning
- ☐ Variabel yang menjadi label/kelas berupa bilangan numeric/angka
- ☐ Algoritma yang dapat digunakan antara lain:
Linear Regression, Neural Network, Support Vector Machine, dsb





Contoh Fungsi Estimasi

Estimasi Performansi CPU

	Cycle time (ns)	Main memory (Kb)		Cache (Kb)	Channels		Performance
	MYCT	MMIN	MMAX	CACH	CHMIN	CHMAX	PRP
1	125	256	6000	256	16	128	198
2	29	8000	32000	32	8	32	269
...							
208	480	512	8000	32	0	0	67
209	480	1000	4000	0	0	0	45

Fungsi Regresi Linear

$$\text{PRP} = -55.9 + 0.0489 \text{ MYCT} + 0.0153 \text{ MMIN} + 0.0056 \text{ MMAX} \\ + 0.6410 \text{ CACH} - 0.2700 \text{ CHMIN} + 1.480 \text{ CHMAX}$$





Regresi Linear

❑ Macam-macam Regresi Linear

Regresi Linear Sederhana

(Single Linear Regression/Bivariat Linear Regression)

- Regresi Linear yang memiliki 1 variabel pemberi pengaruh

Regresi Linear Berganda

(Multiple Linear Regression)

- Regresi Linear yang memiliki lebih dari 1 variabel pemberi pengaruh



Regresi Linear Sederhana

- Digunakan untuk mengukur pengaruh dari variabel prediktor (pemberi pengaruh/independent) terhadap variabel label (terpengaruh/dependent)

- Rumus

Y = variabel terpengaruh

β_0 = konstanta

β_1 = gradien garis

X = variabel pemberi pengaruh

$$Y = \beta_0 + \beta_1 X$$

$$\beta_0 = \bar{Y} - \beta_1 \bar{X}$$

$$\beta_1 = \frac{\sum_{i=1}^n Y_i X_i - \frac{(\sum_{i=1}^n Y_i)(\sum_{i=1}^n X_i)}{n}}{\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}}$$



Regresi Linear Sederhana

Data Transaksi Pengiriman Pesanan Pizza

NO	JARAK	WAKTU
1	0,5	9,95
2	1,1	24,45
3	1,2	31,75
4	5,5	35
5	2,95	25,02
6	2	16,86
7	3,75	14,38
8	0,52	9,6
9	1	24,35
10	3	27,5

NO	JARAK	WAKTU
11	4,12	17,08
12	4	37
13	5	41,95
14	3,6	11,66
15	2,05	21,65
16	4	17,89
17	6	69
18	5,85	10,3
19	5,4	34,93
20	2,5	46,59

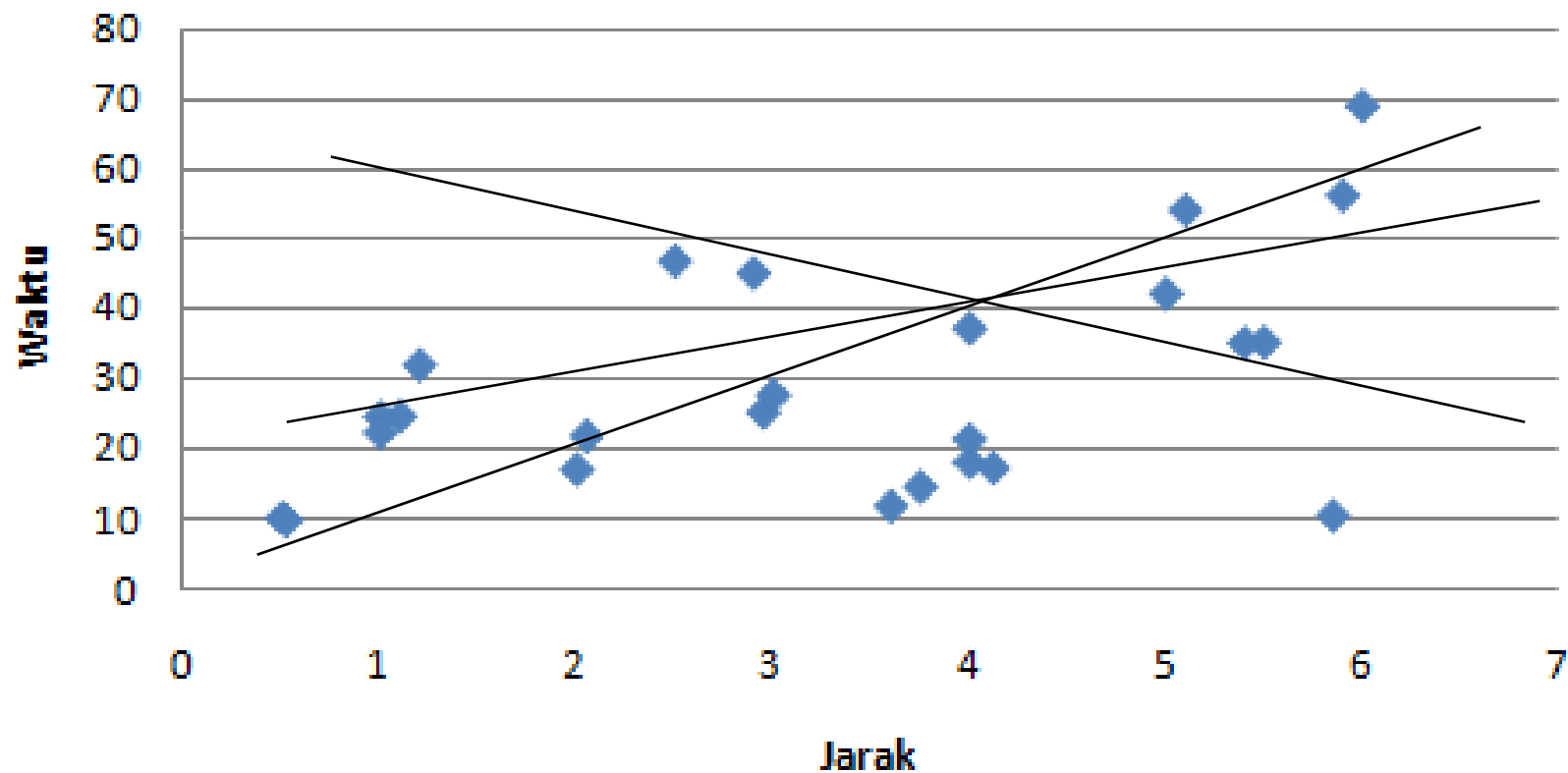
NO	JARAK	WAKTU
21	2,9	44,88
22	5,1	54,12
23	5,9	56,23
24	1	22,13
25	4	21,15





Regresi Linear Sederhana

Transaksi Pengiriman Pesanan Pizza





Regresi Linear Sederhana

no i	jarak x	waktu y	$x_i y_i$	x_i^2
1	0,5	9,95	4,975	0,25
2	1,1	24,45	26,895	1,21
3	1,2	31,75	38,1	1,44
4	5,5	35	192,5	30,25
5	2,95	25,02	73,809	8,7025
6	2	16,86	33,72	4
7	3,75	14,38	53,925	14,0625
8	0,52	9,6	4,992	0,2704
9	1	24,35	24,35	1
10	3	27,5	82,5	9
11	4,12	17,08	70,3696	16,9744
12	4	37	148	16
13	5	41,95	209,75	25
14	3,6	11,66	41,976	12,96
15	2,05	21,65	44,3825	4,2025
16	4	17,89	71,56	16
17	6	69	414	36
18	5,85	10,3	60,255	34,2225
19	5,4	34,93	188,622	29,16
20	2,5	46,59	116,475	6,25
21	2,9	44,88	130,152	8,41
22	5,1	54,12	276,012	26,01
23	5,9	56,23	331,757	34,81
24	1	22,13	22,13	1
25	4	21,15	84,6	16
Total	82,94	725,42	2745,81	353,185
Rata-rata	3,3176	29,0168		

$$\beta_1 = \frac{\sum_{i=1}^n Y_i X_i - \frac{(\sum_{i=1}^n Y_i)(\sum_{i=1}^n X_i)}{n}}{\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}}$$

$$= (2745,81 - ((725,42)(82,94))/25) / (353,185 - ((82,94)^2)/25)$$

$$= 4,35$$

$$\beta_0 = \bar{Y} - \beta_1 \bar{X}$$

$$= 29,0168 - 4,35 * 3,3176$$

$$= 14,59$$

$$Y = \beta_0 + \beta_1 X$$

$$= 14,59 + 4,35 X$$



Regresi Linear Berganda

- ❑ Digunakan untuk mengukur pengaruh antara lebih dari satu variabel prediktor (pemberi pengaruh/independent) terhadap variabel label (terpengaruh/dependent)
- ❑ Rumus $Y = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n$

Y = variabel terpengaruh

β_0 = konstanta

β_1, β_n = gradien garis

x_1, x_n = variabel pemberi pengaruh





Regresi Linear Berganda

Data Transaksi Pengiriman Pesanan Pizza

NO	LAMPU	JARAK	WAKTU
1	2	0,5	9,95
2	8	1,1	24,45
3	11	1,2	31,75
4	10	5,5	35
5	8	2,95	25,02
6	4	2	16,86
7	2	3,75	14,38
8	2	0,52	9,6
9	9	1	24,35
10	8	3	27,5
11	4	4,12	17,08
12	11	4	37
13	12	5	41,95

NO	LAMPU	JARAK	WAKTU
14	2	3,6	11,66
15	4	2,05	21,65
16	4	4	17,89
17	20	6	69
18	1	5,85	10,3
19	10	5,4	34,93
20	15	2,5	46,59
21	15	2,9	44,88
22	16	5,1	54,12
23	17	5,9	56,23
24	6	1	22,13
25	5	4	21,15

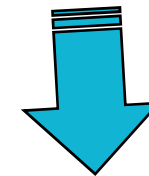
Regresi Linear Berganda

no	lampu	jarak	waktu	x_1^2	x_1x_2	x_1y_i	x_2^2	x_2y_i
i	x1	x2	y					
1	2	0,5	9,95	4	1	19,9	0,25	4,975
2	8	1,1	24,45	64	8,8	195,6	1,21	26,895
3	11	1,2	31,75	121	13,2	349,25	1,44	38,1
4	10	5,5	35	100	55	350	30,25	192,5
5	8	2,95	25,02	64	23,6	200,16	8,7025	73,809
6	4	2	16,86	16	8	67,44	4	33,72
7	2	3,75	14,38	4	7,5	28,76	14,0625	53,925
8	2	0,52	9,6	4	1,04	19,2	0,2704	4,992
9	9	1	24,35	81	9	219,15	1	24,35
10	8	3	27,5	64	24	220	9	82,5
11	4	4,12	17,08	16	16,48	68,32	16,9744	70,3696
12	11	4	37	121	44	407	16	148
13	12	5	41,95	144	60	503,4	25	209,75
14	2	3,6	11,66	4	7,2	23,32	12,96	41,976
15	4	2,05	21,65	16	8,2	86,6	4,2025	44,3825
16	4	4	17,89	16	16	71,56	16	71,56
17	20	6	69	400	120	1380	36	414
18	1	5,85	10,3	1	5,85	10,3	34,2225	60,255
19	10	5,4	34,93	100	54	349,3	29,16	188,622
20	15	2,5	46,59	225	37,5	698,85	6,25	116,475
21	15	2,9	44,88	225	43,5	673,2	8,41	130,152
22	16	5,1	54,12	256	81,6	865,92	26,01	276,012
23	17	5,9	56,23	289	100,3	955,91	34,81	331,757
24	6	1	22,13	36	6	132,78	1	22,13
25	5	4	21,15	25	20	105,75	16	84,6
Total	206	82,94	725,42	2396	771,77	8001,67	353,185	2745,81

$$n\beta_0 + \beta_1 \sum_{i=1}^n x_{i1} + \beta_2 \sum_{i=1}^n x_{i2} = \sum_{i=1}^n y_i$$

$$\beta_0 \sum_{i=1}^n x_{i1} + \beta_1 \sum_{i=1}^n x_{i1}^2 + \beta_2 \sum_{i=1}^n x_{i1}x_{i2} = \sum_{i=1}^n x_{i1}y_i$$

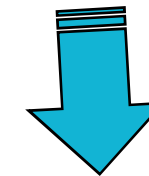
$$\beta_0 \sum_{i=1}^n x_{i2} + \beta_1 \sum_{i=1}^n x_{i1}x_{i2} + \beta_2 \sum_{i=1}^n x_{i2}^2 = \sum_{i=1}^n x_{i2}y_i$$



$$25\beta_0 + \beta_1(206) + \beta_2(82,94) = 725,42$$

$$\beta_0(206) + \beta_1(2396) + \beta_2(771,77) = 8001,67$$

$$\beta_0(82,94) + \beta_1(771,77) + \beta_2(353,18) = 2745,81$$



$$Y = 2,33 + 2,74 X_1 + 1,24 X_2$$



Regresi Linear Berganda

$$25 \beta_0 + 206 \beta_1 + 82,94 \beta_2 = 725,42$$



Persamaan 1

$$206 \beta_0 + 2396 \beta_1 + 771,77 \beta_2 = 8001,67$$



Persamaan 2

$$82,94 \beta_0 + 771,77 \beta_1 + 353,18 \beta_2 = 2745,81$$



Persamaan 3

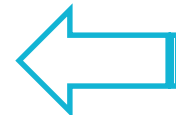
Persamaan 2

$$206 \beta_0 + 2396 \beta_1 + 771,77 \beta_2 = 8001,67 \quad || \times 1 \quad \left. \vphantom{206 \beta_0 + 2396 \beta_1 + 771,77 \beta_2 = 8001,67} \right\} 206 \beta_0 + 2396 \beta_1 + 771,77 \beta_2 = 8001,67$$

Persamaan 1

$$25 \beta_0 + 206 \beta_1 + 82,94 \beta_2 = 725,42 \quad || \times \frac{206}{25} \quad \left. \vphantom{25 \beta_0 + 206 \beta_1 + 82,94 \beta_2 = 725,42} \right\} 206 \beta_0 + 1697,44 \beta_1 + 683,43 \beta_2 = 5977,47$$

Persamaan 4



$$698,56 \beta_1 + 88,34 \beta_2 = 2024,2$$

Persamaan 3

$$82,94 \beta_0 + 771,77 \beta_1 + 353,18 \beta_2 = 2745,81 \quad || \times 1 \quad \left. \vphantom{82,94 \beta_0 + 771,77 \beta_1 + 353,18 \beta_2 = 2745,81} \right\} 82,94 \beta_0 + 771,77 \beta_1 + 353,18 \beta_2 = 2745,81$$

Persamaan 1

$$25 \beta_0 + 206 \beta_1 + 82,94 \beta_2 = 725,42 \quad || \times \frac{82,94}{25} \quad \left. \vphantom{25 \beta_0 + 206 \beta_1 + 82,94 \beta_2 = 725,42} \right\} 82,94 \beta_0 + 683,43 \beta_1 + 275,16 \beta_2 = 2406,65$$

Persamaan 5



$$88,34 \beta_1 + 78,02 \beta_2 = 339,16$$





Regresi Linear Berganda

Persamaan 4

$$698,56 \beta_1 + 88,34 \beta_2 = 2024,2 \quad || \times 1$$

$$\left. \begin{array}{l} 698,56 \beta_1 + 88,34 \beta_2 = 2024,2 \end{array} \right\}$$

Persamaan 5

$$88,34 \beta_1 + 78,02 \beta_2 = 339,16 \quad || \times \frac{88,34}{78,02}$$

$$\left. \begin{array}{l} 100,03 \beta_1 + 88,34 \beta_2 = 384,02 \end{array} \right\} -$$

$$598,53 \beta_1 = 1640,18$$

$$\beta_1 = \frac{1640,18}{598,53} = 2,74$$

Persamaan 5

$$88,34 \beta_1 + 78,02 \beta_2 = 339,16$$

$$88,34 * 2,74 + 78,02 \beta_2 = 339,16$$

$$242,05 + 78,02 \beta_2 = 339,16$$

$$\beta_2 = \frac{339,16 - 242,05}{78,02} = 1,24$$





Regresi Linear Berganda

Persamaan 1 $25 \beta_0 + 206 * 2,74 + 82,94 * 1,24 = 725,42$

$$25 \beta_0 + 564,44 + 102,85 = 725,42$$

$$\beta_0 = \frac{725,42 - 667,29}{25}$$

$$\beta_0 = 2,33$$

$$Y = 2,33 + 2,74 X_1 + 1,24 X_2$$



A collage of three images is positioned on the left side of the slide. The top image shows several gold bars with serial numbers like "100153" and "100154". The middle image shows a hand holding a blue pen, writing on a document that has a line graph and the number "+12.7%". The bottom image shows a close-up of a blue line graph on a black background with the number "2" visible.

Terima Kasih