

CHAPTER 4

ANALYSIS AND DESIGN

1.5. Analysis

This research is Simple Linear Regression, used in Data Mining to predict all linear data against a value. In Simple Linear Regression, data is modeled in the form of a graph in the form of two-dimensional lines, so it takes the variables X and Y. And in this study using test data as many as 1000 datasets, with X as a free variable and Y as a bound variable.

In Linear Regression, the variable Y is referred to as *the response variable* while X is referred to as *the predictor variable*. The two variables are formulated statically with the formula $y = \alpha + \beta x$.

The value of y in the above formulation is considered to be the value of constant, while the value α and β is *the regression coefficient* which affects the delineation of data in the two-dimensional graph.

The value α and β can be searched using the *least square* method which functions to minimize the error value between the actual data and the data of the predicate result. Given the sample value of the data S with dots $(x_1, y_1), (x_2, y_2), \dots (x_3, y_3)$, then the *regression coefficient* can be searched using the following formula :

Constant(α) :

$$\alpha = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

Coefficient (β)

$$\beta = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

where \bar{x} is the average of x_1, x_2, \dots, x_n and \bar{y} are the averages of y_1, y_2, \dots, y_n .

1.6. Design

This is the X and Y value data for the 1000 datasets already provided.

Table 1 Values of Variable X and Variable Y

X	Y
77	79,77515201
21	23,17727887
22	25,60926156
20	17,85738813
36	41,84986439
15	9,805234876
62	58,87465933
95	97,61793701
20	18,39512747
5	8,746747654
4	2.811415826
...	...

The above score data can be depicted in the form of a two-dimensional linear graph where the mid and final exam score data are in the form of linear lines. The x point is the data for the variable X while the y point is the data of the Variable Y.