**Algoritma Normalisasi**

Ketika sebuah dataset akan digunakan sebagai acuan untuk pembuatan model, ada beberapa standar prosedur harus disiapkan. Salah satunya adalah dataset tersebut harus dilakukan normalisasi atau standarisasi agar setiap variable pada dataset memiliki proporsi yang sama. Proses normalisasi dibutuhkan karena rentang nilai pada sebuah dataset bervariasi dan tidak selalu sama, di dalam machine learning, sebuah fungsi tujuan tidak akan berfungsi dengan baik jika sebuah dataset digunakan tanpa melalui normalisai data.

1. Normalisasi data dengan menggunakan Min-Max Scalling

Min-Max scalling merupakan salah satu metode sederhana dengan melakukan *rescalling* pada berbagai variable / feature untuk rentang skala di [0,1] atau [-1,1].

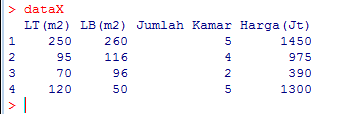
Rumus umum Min-Max Scalling:

http://sebastianraschka.com/Images_old/2014_about_feature_scaling/2014_about_feature_scaling_2.png

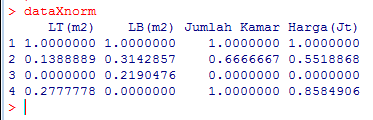
Keterangan :

* Xnorm = data ternormalisasi
* X = data sebelum ternormalisasi
* Xmin = data paling minimum
* Xmax = data paling maksimum

Contoh Kasus: terdapat sebuah data rumah sebagai berikut



Dari dataset tersebut akan dilakukan normalisasi denganmenggunakan Min-Max Scalling:



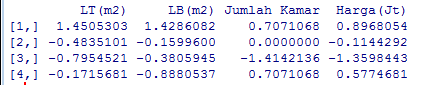
Data Hasil Normalisasi :

1. Z-Score Normalization

Hasil dari standardisasi (atau Z-score normalisasi) adalah bahwa fitur akan terskala sehingga mereka akan memiliki sifat-sifat dari distribusi normal standar dengan μ = 0 and σ = 1, di mana μ adalah mean (rata-rata) dan σ adalah standar deviasi dari mean; nilai standar (juga disebut skor z) dari sampel dihitung sebagai berikut:

http://sebastianraschka.com/Images_old/2014_about_feature_scaling/2014_about_feature_scaling_1.png Keterangan :

* + - Z = Standarized data
    - X = Data
    - μ = mean
    - σ = stdev

1. Sigmoidal normalization  
   Melakukan normalisasi datasecara nonlinier ke dalam range -1 - 1 dengan menggunakan fungsi sigmoid.  
      
   Rumus: Soft Computation Research Group, EEPISITS  
    Rumus:  
   newdata = (1-e^(-x))/(1+ e^(-x))  
   dimana:  
   x = (data-mean)/std  
   e = nilai eksponensial (2,718281828)  
    Metode ini sangat berguna pada saat data-data yang  
   ada melibatkan data-data outlier.

**Algoritma Denormalisasi**

Denormalisasi adalah mengembalikan ukuran data yang telah dinormalisasi sebelumnya untuk mendapatkan data yang asli. Denormalisasi dilakukan pada hasil keluaran dari pelatihan berupa prediksi curah hujan. Adapun rumus dari denormalisasi yaitu sebagai berikut :

1. MinMax Scalling  
     
   Denormalisasi = Y \* (max-min) + min  
     
   Y = hasil keluaran dari pelatihan  
   Min = data minimum  
   Max = data maximum
2. Mean Scalling  
     
   Denormalisasi = Y \* (STD + SM)  
     
   Y = hasil keluaran dari pelatihan  
   STD = sample standard deviasi  
   SM = sample mean