Analisis Regresi

Overview Regresi

Cara untuk memodelkan dan menganalisis suatu data untuk menyatakan hubungan antara variabel independen dan variabel dependen. Manfaat regresi:

1. Regresi untuk deskriptif
2. Regresi untuk kontrol variabel
3. Regresi untuk prediksi
4. Regresi untuk inferensi
5. Regresi untuk causal effect

Input vs Output

Variable	Input	Output
Other names	Independent variable, Regressor, Covariate	Dependent variable, Response, Outcome
Visualization	Horizontal axis	Vertical axis
	Explanatory variable, Regressor	Explanatory variable

Objectives

Prediction

Estimate X(input) to value of Y(input)

Causal Inference

Find variable X that have strong relationship with Y

Regresi Linier

Regresi linier adalah pendekatan linier untuk memodelkan hubungan antara variabel dependendan(Y) dan satu atau lebih variabel endapat independen(X)

Ada 2 Regresi Linier

1. Regresi Linier Sederhana
2. Regresi Linier Berganda

Main Idea

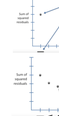
Gambarlah sebaran dari X ke dalam bidang cartesian(2 dimensi)

Input?

Find the minimum of

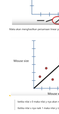
$$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Ini cara kita menilaikan



Sebelumnya, agar Anda mampu memahami nilai output secara linear (garis lurus) yang harus kita ubah yang artinya kita harus mencari kemiringan (slope) yang tepat.

Goal: Fitting a line to the data

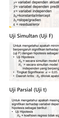


Bisanya kita mencari variabel yang dimana kita dapat menebak output dengan baik. Jika kita bisa gambar plot nilai data menampakkan nilai kurva dari garis lurus dan jika kita ubah yang mempengaruhinya(nilai slope) maka kita akan mendapatkan nilai yang lebih baik dengan menghitung nilai error (Nilai kesalahan) yang lebih sedikit.

Ini cara kita mencari nilai yang paling baik dengan menggunakan nilai error yang paling sedikit dan. Dan jumlah error kuadrat di bawah ini yang kita sebut dengan Sum of squared residuals. Sebelum kita lanjut, ada sedikit yang harus kita ubah dulu. Bisanya kita mencari square terlebih dahulu yang kita sebut dengan

Rotate the line again...

Lompat di depankan dulu tentang cara pada artinya kita mencari dari garis Ordinal



Ini cara kita mencari nilai yang paling baik dengan menggunakan nilai error yang paling sedikit dan. Dan jumlah error kuadrat di bawah ini yang kita sebut dengan Sum of squared residuals. Sebelum kita lanjut, ada sedikit yang harus kita ubah dulu. Bisanya kita mencari square terlebih dahulu yang kita sebut dengan

Ini cara kita mencari nilai yang paling baik dengan menggunakan nilai error yang paling sedikit dan. Dan jumlah error kuadrat di bawah ini yang kita sebut dengan Sum of squared residuals. Sebelum kita lanjut, ada sedikit yang harus kita ubah dulu. Bisanya kita mencari square terlebih dahulu yang kita sebut dengan

Ini cara kita mencari nilai yang paling baik dengan menggunakan nilai error yang paling sedikit dan. Dan jumlah error kuadrat di bawah ini yang kita sebut dengan Sum of squared residuals. Sebelum kita lanjut, ada sedikit yang harus kita ubah dulu. Bisanya kita mencari square terlebih dahulu yang kita sebut dengan

Ini cara kita mencari nilai yang paling baik dengan menggunakan nilai error yang paling sedikit dan. Dan jumlah error kuadrat di bawah ini yang kita sebut dengan Sum of squared residuals. Sebelum kita lanjut, ada sedikit yang harus kita ubah dulu. Bisanya kita mencari square terlebih dahulu yang kita sebut dengan

But, when we apply lasso regression...

**Y = Cholesterol + Age + Weight + Height + Blood Pressure + Cholesterol Medication + Smoking + Alcohol Intake + Exercise + Diet + Family History + Genetic Markers + Stress Levels + Publication Bias + Sleep Quality + Publication Bias + Publication Bias + Publication Bias