

Nama : Indira Septianita Larasati

NIM :H071191023

Tugas : Regresi Linear dan Logistik

Linear Regression,

Linear Regression (Regresi Linear) adalah suatu regresi linear yang digunakan untuk mengestimasi atau memprediksi hubungan antara dua variabel dalam penelitian kuantitatif. Dimana regresi linear ini mampu membuat satu asumsi tambahan yang mengkorelasikan antara variabel independen dan dependen melalui garis yang paling sesuai dari titik data garis lurus. Artinya, bukan kurva atau semacam faktor penelompokkan.

Meski demikian, regresi linear memiliki keterbatasan, karena dalam data terbaik pun tidak menjelaskan dan menampilkan data yang lengkap. Analisis regresi biasanya digunakan dalam penelitian untuk menetapkan bahwa ada korelasi antar variabel. Namun, korelasi yang tidak sama sebab akibat adanya hubungan antara dua variabel tidak berarti yang satu menyebabkan yang lainnya terjadi. Bahkan garis dalam regresi linear sederhana yang cocok dengan titik data mungkin tidak menjamin hubungan sebab-akibat.

Kelebihan linear regression,

Ketika hubungan antara variabel independen dan dependen memiliki hubungan linear, algoritma ini adalah yang terbaik untuk digunakan, karena ini adalah yang paling kompleks dibanding algoritma lain yang juga menemukan hubungan antara variabel independen dan dependen. Metode ini mampu digunakan untuk memprediksi nilai yang ada pada masa depan. Hal ini sejalan dengan fungsi dari analisis regresi yang dapat digunakan untuk peramalan dan prediksi.

Kekurangan dari linear regression,

Pada kenyataannya, dalam data real, jarang masalah didunia yang menunjukkan hubungan yang jelas antara variabel dependen dan independen. Hal ini dapat membuat model yang tidak cukup bagus, disebabkan karena kesalahan dalam memilih variabel yang digunakan untuk analisis.

Dalam regresi linier, hasilnya (variabel dependen) kontinu. Itu dapat memiliki salah satu dari sejumlah kemungkinan nilai yang tak terbatas. Sedangkan dalam regresi logistik, hasil (variabel dependen) hanya memiliki sejumlah nilai yang mungkin.

- **Variabel dependen**

Regresi logistik digunakan ketika variabel respon bersifat kategoris. Misalnya, ya/tidak, benar/salah, merah/hijau/biru, 1/2/3/4, dan lain lain.

Regresi linier digunakan ketika variabel respons Anda kontinu. Misalnya, berat, tinggi, jumlah jam, dan sebagainya.

- **Persamaan**

Regresi linier memberikan persamaan yang berbentuk $Y = mX + C$, artinya persamaan dengan derajat 1. Sedangkan regresi logistik memberikan persamaan yang berbentuk $Y = \frac{e^X}{e^X + e^{-X}}$

- **Koefisien interpretasi**

Dalam regresi linier, interpretasi koefisien variabel independen cukup mudah (mis. Memegang semua variabel lain konstan, dengan peningkatan unit dalam variabel ini, variabel dependen diharapkan meningkat/berkurang sebesar xxx).

Namun, dalam regresi logistik, tergantung pada keluarga (binomial, Poisson, dll.) Dan tautan (log, logit, inverse-log, dll.) Yang Anda gunakan, interpretasinya berbeda.

- **Teknik minimalisasi kesalahan**

Regresi linier menggunakan *metode kuadrat terkecil* untuk meminimalkan kesalahan dan mencapai kesesuaian terbaik, sedangkan regresi logistik menggunakan *metode kemungkinan maksimum* untuk sampai pada solusi.

Regresi linier biasanya diselesaikan dengan meminimalkan kesalahan kuadrat terkecil dari model terhadap data, oleh karena itu kesalahan besar akan dihukum secara kuadratik.

Regresi logistik justru sebaliknya. Menggunakan fungsi kerugian logistik menyebabkan kesalahan besar akan dihukum konstan asimptotik.

Pertimbangkan regresi linier pada hasil kategori $\{0, 1\}$ untuk melihat mengapa ini merupakan masalah. Jika model Anda memprediksi hasilnya adalah 38, ketika kebenarannya 1, Anda tidak kehilangan apa-apa. Regresi linier akan mencoba mengurangi 38 itu, logistik tidak akan (sebanyak)

Dalam regresi linier, hasilnya (variabel dependen) kontinu. Itu dapat memiliki salah satu dari sejumlah kemungkinan nilai yang tak terbatas. Dalam regresi logistik, hasil (variabel dependen) hanya memiliki sejumlah nilai yang mungkin.

Misalnya, jika X berisi area dalam kaki persegi rumah, dan Y berisi harga jual yang sesuai dari rumah-rumah itu, Anda bisa menggunakan regresi linier untuk memprediksi harga jual sebagai fungsi ukuran rumah. Sementara kemungkinan harga jual mungkin tidak benar-benar *ada*, ada begitu banyak nilai yang mungkin bahwa model regresi linier akan dipilih.

Sebaliknya, jika kita ingin memprediksi, berdasarkan ukuran, apakah sebuah rumah akan dijual lebih dari \$ 200 ribu, maka kita akan menggunakan regresi logistik. Output yang mungkin adalah Ya, rumah akan menjual lebih dari \$ 200 ribu, atau tidak, rumah tidak akan.