

Classification 1

- ***Classification***

Classification adalah metode yang digunakan untuk memprediksi **target variable** bertipe kategorik (**factor**).

- ***Probability***

Kemungkinan terjadinya suatu kejadian.

- ***Odds***

Ukuran yang dapat menjelaskan probability. Dimana odds bisa di dapatkan dari $p/(1-p)$ dimana p adalah peluang suatu kejadian terjadi.

- ***Odds ratio***

Odds Ratio adalah perbandingan antara dua odds.

- ***Sigmoid function***

Sigmoid function merupakan fungsi yang digunakan untuk mentransformasi nilai prediksi ke nilai peluang yaitu antara 0 hingga 1.

- ***Standarization***

Proses untuk menyeragamkan skala data yang berbeda.

- ***Class imbalance***

Keadaan dimana jumlah observasi antar kelas tidak seimbang.

- ***Cross Validation***

Proses untuk membagi data menjadi data train dan data test.

- ***Data train***

Data yang digunakan untuk membuat model.

- ***Data test***

Data yang digunakan untuk menguji kebaikan model.

- ***Overfitting***

Keadaan dimana model yang dibuat hanya dapat memprediksi dengan baik **data train**. Namun, ketika melakukan prediksi pada **data test**, model tersebut tidak dapat memprediksi dengan baik.

- ***Independence of observations***

Antar observasi independen satu sama lain.

- ***Null deviance***

Null deviance menunjukkan seberapa baik **target variable** diprediksi oleh model berdasarkan nilai intercept.

- ***Residual deviance***

Residual deviance menunjukkan seberapa baik **target variable** diprediksi oleh model berdasarkan intercept dan semua **predictor** yang digunakan.

- ***Maximum likelihood estimator***

Maximum likelihood estimator merupakan pendekatan statistik untuk memperkirakan parameter pada model **logistic regression**.

- ***Perfect Separation***

Sebuah kondisi dimana ada 1 variabel yang dapat memisahkan kelas target secara sempurna.