

Neural Network

- ***Target variable***

Variable yang nilainya dipengaruhi oleh *predictor variable* dan akan diprediksi nilainya, sering disebut sebagai respon/dependent variable.

- ***Predictor variable***

Variable yang mempengaruhi dan digunakan untuk memprediksi nilai *target variable*, sering disebut sebagai independent variable.

- ***Feature engineering***

Menambahkan informasi (variabel/kolom) berdasarkan informasi dari variabel lain yang sudah ada.

- ***Missing value***

Keadaan di mana data memiliki nilai yang hilang (tidak diketahui nilainya/NA).

- ***Standarization***

Proses untuk menyeragamkan interval data yang berbeda, sering disebut sebagai *scaling* atau *normalization*.

- ***Data train***

Bagian data yang digunakan untuk membuat model (training model).

- ***Data test***

Bagian data yang digunakan untuk mengevaluasi performa model (testing model).

- ***Cross Validation***

Proses membagi data lengkap/utuh menjadi dua bagian, yaitu *data train* dan *data test*.

- ***Class imbalance***

Keadaan di mana suatu kategori pada *target variable* memiliki proporsi lebih mendominasi dibandingkan kategori lainnya.

- ***Sampling***

Proses mengambil sebanyak n bagian data, umumnya dilakukan secara acak.

- ***Down-sample***

Proses *sampling* pada observasi kelas mayoritas, sehingga jumlahnya sama dengan jumlah observasi pada kelas minoritas.

- ***Up-sample***

Proses *sampling* pada observasi kelas minoritas, sehingga jumlahnya sama dengan jumlah observasi pada kelas mayoritas.

- ***Optimization***

Proses mengoptimalkan suatu nilai. Pada model neural network bertujuan untuk meminimumkan fungsi error dengan menggunakan fungsi turunan.

- ***Nodes***

Unit terkecil pada arsitektur neural network yang berfungsi untuk menyimpan informasi dan meneruskan informasi tersebut, sering disebut sebagai *neuron*.

- ***Input layer***

Lapisan pertama pada arsitektur neural network yang terdiri dari kumpulan *nodes*. Jumlah *nodes* pada *input layer* bergantung pada jumlah *predictor*.

- ***Output layer***

Lapisan terakhir pada arsitektur neural network yang terdiri dari satu atau beberapa *nodes*, bergantung pada tujuannya (regresi/klasifikasi).

- ***Hidden layer***

Lapisan yang terletak di antara *input layer* dan *output layer*. Jumlah *hidden layer* dan jumlah *nodes* di setiap *hidden layer* ditentukan oleh *user*.

- ***Weight***

Bobot yang menggambarkan besar informasi yang diteruskan dari setiap *nodes*. *Weight* diinisialisasi secara acak.

- ***Bias***

Pada *linear regression* sama seperti nilai *intercept* (b_0).

- ***Activation function***

Fungsi yang digunakan untuk mengubah interval nilai informasi yang masuk ke setiap *nodes* pada *hidden layer* dan *output layer*.

- ***Cost function***

Fungsi error yang akan dioptimisasi pada model neural network.

- ***Feedforward***

Proses pada neural network yang dimulai dari *input layer* hingga menghasilkan nilai prediksi pada *output layer*.

- ***Backpropagation***

Proses pada neural network ketika melakukan *optimization* dari cost function dan melakukan update *weight*.

- ***Gradient***

Hasil turunan pertama dari fungsi error/*cost function* yang merupakan kemiringan dari titik potong *cost function* di titik tersebut.

- ***Dummy Variable***

Hasil transformasi variabel kategorik dengan nilai 0 atau 1. Variabel ini digunakan untuk membuat data kategorik yang bersifat kualitatif menjadi kuantitatif.

- ***Learning rate***

Besar nilai yang menentukan seberapa cepat proses update *weight*.

- ***Batch***

Proses membagi data menjadi beberapa bagian. Setiap bagian akan diikutsertakan ke dalam proses *feedforward* dan *backpropagation*.

- ***Epoch***

Keadaan di mana seluruh *batch* telah menyelesaikan proses *feedforward* dan *backpropagation*.

- ***Confusion Matrix***

Metriks yang digunakan untuk mengukur kebaikan model classification, terdiri dari *accuracy*, *recall*, *specificity*, dan *precision*.

- ***Sum squared error (SSE)***

Jumlah dari error kuadrat. Ukuran yang bisa digunakan untuk mengukur kebaikan model.

- *Mean squared error (MSE)*

Rata-rata dari error kuadrat. Ukuran yang bisa digunakan untuk mengukur kebaikan model.

- *Root Mean squared error (RMSE)*

Akar kuadrat dari *MSE*.

- *Mean absolute error (MAE)*

Rata-rata dari absolut error. Ukuran yang bisa digunakan untuk mengukur kebaikan model.