

Daftar Istilah

Istilah	Deskripsi	Halaman pertama kali digunakan
Laten	: tersembunyi; terpendam; tidak kelihatan	6
<i>Wordcloud</i>	: representasi visual dari kumpulan kata-kata dalam sebuah teks. Ini adalah metode yang populer untuk menggambarkan frekuensi kata dalam teks dengan cara yang lebih menarik dan mudah dimengerti secara visual.	7
<i>Overfitting</i>	: merupakan fenomena di mana data pelatihan terlalu cocok yang menimbulkan risiko menyesuaikan model dengan <i>noise</i> dalam data, di mana dengan spesifik mengingat keanehan-keanehan yang ada dalam data pelatihan daripada menemukan pola prediksi yang lebih umum dan tidak dapat mengenali data baru dengan baik	9
Akuisisi	: pemindahan kepemilikan perusahaan atau aset (dalam industri perbankan terjadi apabila pembelian saham di atas 50%), pengambilalihan kepemilikan perusahaan atau asset.	17
Akumulasi	: tambahan periodik suatu dana dari bunga atau tambahan lain pada tambahan laba neto pada laba yang ditahan	17
Semantik	: ilmu tentang makna kata dan kalimat; pengetahuan mengenai seluk-beluk dan pergeseran arti kata	22
Leksikon	: daftar istilah dalam suatu bidang disusun menurut abjad dan dilengkapi dengan keterangannya	22
Korpus	: sumber daya bahasa yang terdiri dari kumpulan teks yang besar.	27

Alfanumerik	:	istilah yang mengacu pada kombinasi huruf dan angka dalam suatu teks atau karakter	27
divergensi	:	konsep dalam matematika dan statistika yang mengukur seberapa jauh dua distribusi atau fungsi probabilitas berbeda satu sama lain.	29
<i>Library</i> Python	:	kumpulan modul dan fungsi yang telah dibangun sebelumnya yang dapat digunakan untuk memperluas kemampuan bahasa pemrograman Python.	35
Koheren	:	berhubungan atau bersangkutan paut	42
Data diskrit	:	jenis data yang terdiri dari nilai-nilai terpisah atau terbatas pada himpunan nilai yang terdefinisi dengan jelas.	47
<i>throughput</i>	:	jumlah dokumen atau teks yang dapat diproses atau dianalisis dalam suatu periode waktu tertentu	47
Bias	:	Bias pada machine learning mengacu pada kesalahan sistem yang disebabkan oleh penyederhanaan yang berlebihan dalam model atau algoritma.	54
<i>Variance</i>	:	istilah yang digunakan dalam statistik dan probabilitas untuk mengukur sejauh mana data poin tersebar di sekitar nilai rata-rata atau <i>mean</i>	82
<i>Heatmap</i>	:	bentuk visualisasi data yang digunakan untuk menggambarkan hubungan antara dua set data dengan menggunakan skala warna.	97