

DAFTAR ISTILAH

Istilah	Deskripsi	Halaman
Algoritma <i>K-Means</i>	: Algoritma untuk <i>cluster</i> n objek berdasarkan atribut menjadi k partisi, dimana $k < n$. Gambar berikut ini menunjukkan <i>k-means clustering</i> algoritma dalam tindakan, untuk kasus dua dimensi	v
<i>Business Understanding</i>	: Adalah aktifitas untuk memahami tujuan bisnis dan kebutuhan proyek.	xi
<i>Centroid</i>	: Titik pusat dari setiap kelompok dalam algoritma <i>K-Means</i> , mewakili nilai rata-rata dari anggota kelompok tersebut.	xi
<i>Clustering</i>	: Sebuah metode <i>Unsupervised Learning</i> yang bertujuan untuk melakukan pengelompokan data berdasarkan kemiripan/jarak antar data. <i>Clustering</i> memiliki karakteristik dimana anggota dalam satu cluster memiliki kemiripan yang sama atau jarak yang sangat dekat, sementara anggota antar cluster memiliki kemiripan yang sangat berbeda atau jarak yang sangat jauh	i
<i>Cluster</i>	: Langkah-langkah persiapan data sebelum diterapkan pada algoritma <i>K-Means</i> , seperti normalisasi, penghapusan <i>missing values</i> , dan transformasi data	ix
<i>Cleaning</i>	: Proses menghilangkan data yang tidak sesuai atau tidak lengkap dari data yang terkumpul.	25
Corona Virus Disease (COVID-19)	: Penyakit baru yang dapat menyebabkan gangguan pernapasan dan radang paru. Penyakit ini disebabkan	xi

oleh infeksi *Severe Acute Respiratory Syndrome Coronavirus 2* (SARS-CoV-2).

<i>Cross-Industry Standard Process for Data Mining (CRISP-DM)</i>	:	Merupakan sebuah metodologi yang digunakan dalam proyek data mining dan analisis data	xi
<i>DataFrame</i>	:	Adalah struktur data tabular (tabel) dua dimensi yang memiliki baris dan kolom.	ix
<i>Data Understanding</i>	:	Adalah tahapan untuk mengumpulkan dan memahami data yang akan digunakan.	xi
<i>Data Preparation</i>	:	Adalah tahapan untuk membersihkan, mengubah, dan memformat data untuk analisis.	xi
<i>Dataset</i>	:	kumpulan informasi yang terkumpul dari berbagai sumber, yang akan digunakan untuk tujuan analisis	viii
<i>Deployment</i>	:	Adalah tahapan untuk mengimplementasikan hasil proyek ke dalam operasional bisnis.	xii
<i>Droplet</i>	:	Cairan atau cipratan liur yang dikeluarkan seseorang dari hidung atau mulut saat bersin, batuk, bahkan berbicara	xii
<i>Elbow Method</i>	:	Sebuah metode yang digunakan dalam analisis statistik untuk menentukan jumlah klaster (<i>cluster</i>) yang optimal dalam data. Metode ini dinamakan " <i>elbow</i> " karena bentuk grafik yang dihasilkan menyerupai siku (<i>elbow</i>).	ix
<i>Evaluation</i>	:	Adalah tahapan untuk mengevaluasi hasil model dan memastikan bahwa tujuan bisnis tercapai.	xii

Faktor Risiko Tingkat kematian	:	Hal-hal yang berdampak pada peningkatan atau penurunan tingkat kematian (jumlah kematian dalam suatu populasi selama periode waktu tertentu).	i
<i>Klasterisasi</i>	:	Adalah proses pengelompokan data menjadi beberapa kelompok atau cluster sehingga dalam satu kelompok memiliki kemiripan yang maksimum dan data antar kelompok memiliki kemiripan yang minimum.	xiii
MERS	:	<i>Middle East Respiratory Syndrome</i>	xiii
<i>Pattern</i>	:	Pola yang terdeteksi di dalam data yang terkumpul	i
<i>Preprocessing Data</i>	:	Karakteristik atau variabel yang digunakan untuk mengelompokkan data ke dalam <i>cluster</i> .	xiii
SARS	:	<i>Severe Acute Respiratory Syndrome</i>	xii
<i>Visualization</i>	:	Proses menampilkan data dalam bentuk visual seperti grafik atau diagram untuk memudahkan pemahaman dan interpretasi data.	xii
<i>Visualization Clustering</i>	:	Representasi grafis dari hasil clustering, seperti peta warna atau diagram pencar, untuk memahami pola dan distribusi kelompok.	xii