

DAFTAR ISTILAH

<i>Bag of Words</i>	: Metode untuk mengubah data teks menjadi vektor dengan menghitung frekuensi kemunculan kata pada seluruh dokumen.
BPS	: Badan Pusat Statistik, merupakan lembaga statistik di Indonesia.
<i>Consumercentric</i>	: Strategi bisnis yang menempatkan konsumen sebagai prioritas utama berdasarkan kebutuhan dan minat konsumen.
<i>Cosine Similarity</i>	: Mengukur kesamaan antara dua vektor ruang produk dalam.
<i>Database</i>	: Kumpulan data yang saling terorganisir dan saling terhubung sehingga dapat diakses dengan mudah.
<i>Dataset</i>	: Kumpulan data mentah berupa tabel yang dapat diolah lebih lanjut.
<i>Data Cleaning</i>	: Proses menghilangkan dan memperbaiki kesalahan, ketidakkonsistenan, atau informasi yang tidak relevan dari data sebelum analisis atau pemodelan.
<i>Data Filtering</i>	: Proses menyaring data maupun menghapus data berdasarkan fitur tertentu dari <i>dataset</i> .
<i>Data Mining</i>	: Proses pengumpulan dan pengolahan data yang tersembunyi dengan memanfaatkan eksplorasi <i>database</i> .
<i>Data Preprocessing</i>	: Tahap awal pengolahan data mentah agar siap digunakan.
<i>Encoding</i>	: Mengubah bentuk data agar dapat dibaca oleh sistem.
<i>Gibbs Sampling</i>	: Metode yang digunakan untuk mendapatkan estimasi parameter model dengan cara membangkitkan parameter model melalui iterasi.

KDD	: <i>Knowledge Discovery Database</i> , merupakan salah satu kerangka kerja pada <i>data mining</i> .
LDA	: <i>Latent Dirichlet Allocation</i> , merupakan metode <i>unsupervised learning</i> pada pemodelan topik dengan dasar teorema Bayes.
<i>Library</i>	: Koleksi pustaka program yang dapat digunakan dalam pemrograman.
LSA	: <i>Latent Semantic Analysis</i> , merupakan merupakan metode <i>unsupervised learning</i> pada pemodelan topik dengan dasar perhitungan matriks.
NLP	: <i>Natural Language Processing</i> , merupakan kecerdasan buatan yang mengolah bahasa alami.
<i>Node</i>	: Simpul yang berperan untuk menerima, meneruskan, atau menyebarkan data.
NPMI	: <i>Normalized Pointwise Mutual Information</i> , merupakan nilai normalisasi dari probabilitas dua peristiwa yang terjadi secara bersamaan.
PDB	: Produk Domestik Bruto, merupakan total nilai produksi dan jasa yang dihasilkan semua orang atau perusahaan dalam satu negara.
PLSA	: <i>Probabilistic Latent Semantic Analysis</i> , merupakan metode <i>unsupervised learning</i> pada pemodelan topik dengan dasar perhitungan probabilitas distribusi multinomial.
PSBB	: Pembatasan Sosial Berskala Besar.
<i>Set</i>	: Himpunan.
<i>Sliding Window</i>	: Teknik komputasi yang bertujuan untuk mengurangi penggunaan <i>loop</i> bersarang dan menggantinya dengan <i>loop</i> tunggal, sehingga mengurangi kompleksitas waktu.
<i>Stemming</i>	: Proses penguraian bentuk dari suatu kata menjadi bentuk kata dasarnya.

<i>Stopword</i>	: Kata umum yang biasanya muncul dalam jumlah besar dan dianggap tidak memiliki makna.
<i>Subset</i>	: Himpunan bagian.
<i>Text Mining</i>	: Analisis teks tidak terstruktur dalam skala besar.
<i>Tokenization</i>	: Pemisahan dokumen teks yang dapat berupa (frasa, kalimat, atau paragraf) menjadi unit yang lebih kecil seperti kata tunggal atau istilah.
TF-IDF	: <i>Term Frequency-Inverse Document</i> , merupakan metode statistik untuk mengukur pentingnya kata dalam dokumen atau kumpulan dokumen.
<i>Web Scraping</i>	: Teknik yang digunakan untuk mendapatkan informasi-informasi yang terdapat dalam situs.
WEF	: World Economic Forum, merupakan organisasi internasional independen yang berkomitmen untuk meningkatkan keadaan perekonomian dunia.