

## DAFTAR ISTILAH

- Analisis Sentimen* : Studi komputasi tentang pendapat, sentimen, emosi, penilaian, dan sikap terhadap suatu entitas, seperti produk, layanan, organisasi, individu, masalah, acara, topik, dan atributnya.
- Knowledge Discovery in Databases* : Sebuah model proses konseptual yang menjelaskan teori-teori komputasi dan alat-alat yang digunakan untuk mendukung ekstraksi informasi dari data
- Text Mining* : Proses pengetahuan intensif dimana pengguna berinteraksi dengan dokumen-dokumen dari waktu ke waktu dengan menggunakan alat analisis.
- Text Preprocessing* : Mempersiapkan dokumen teks yang tidak terstruktur menjadi data terstruktur yang dapat digunakan untuk langkah selanjutnya dengan cara menghilangkan data yang salah atau *noise*, menyamakan bentuk kata dan meminimalisir volume kata.
- Cleaning* : Proses menghapus angka, tanda baca, dan simbol yang ditemukan pada teks atau kalimat.
- Case Folding* : Proses mengubah setiap huruf pada data yang digunakan menjadi huruf kecil.
- Tokenization* : Proses penyeragaman untuk mengubah data yang tidak baku menjadi bahasa baku.
- Stemming* : Proses pemetaan kata menjadi bentuk dasarnya.
- Term Weighting* : Proses pemberian nilai atau bobot terhadap sebuah istilah dalam suatu parameter, sebelum disimpan pada sistem.
- Term Frequency* : Frekuensi suatu kata muncul dalam teks.
- Inverse Document Frequency* : Nilai invers dari teks yang memiliki kata tersebut.
- Naïve Bayes Classifier* : Metode yang digunakan untuk mengkategorikan atau melihat pendapat atau kecenderungan opini terhadap

suatu masalah atau objek, apakah hal tersebut berkategori positif atau negatif.

- K-Fold Cross Validation* : Metode melakukan perulangan sebanyak nilai 'k', dimana data akan dimasukkan secara acak sehingga sistem akan melakukan uji dengan menggunakan data input acak. Dilakukan untuk menghapus bias pada data.
- Confusion Matrix* : Matriks persegi dimana baris mewakili sebenarnya dari *instance* dan kolom mewakili kelas prediksi.