DAFTAR ISTILAH

Argumen	:	Nomina / frasa nominal yang bersama-sama dengan predikat membentuk preposisi.
Constituent	:	Sebuah kata hingga kelompok kata yang berfungsi sebagai satu kesatuan unit dalam struktur hirarki kalimat.
Data testing	:	Data/vektor yang belum diketahui / dianggap belum diketahui kelasnya untuk selanjutnya diprediksi kelasnya menggunakan model klasifikasi yang sudah dibangun.
Data training	:	Data/vektor yang sebelumya sudah diketahui kelasnya, dan digunakan untuk membangun model klasifikasi.
Dataset	:	Kumpulan data.
Feature	:	Kuantitas yang digunakan sebagai dasar ukuran dalam menilai suatu objek.
File XML	:	Bahasa *markup* untuk keperluan umum yang disarankan oleh W3C untuk membuat dokumen *markup* guna keperluan pertukaran data antar sistem yang beraneka ragam.
F-score	:	Rata-rata harmonik berbobot dari *precision* dan *recall*. Bisa juga disebut *f-meassure*.
Hyperplane	:	Garis batas yang memisahkan antar dua kelas.
Klasifikasi	:	Suatu pekerjaan untuk melakukan pembelajaran terhadap fungsi target f yang memetakan setiap vektor (*set* fitur) x ke dalam satu dari sejumlah label kelas y yang tersedia. Pekerjaan pelatihan tersebut akan menghasilkan suatu model yang kemudian disimpan sebagai memori.
Korpus	:	Koleksi dokumen atau data berupa teks yang digunakan dalam studi kasus *text mining.*
Parse Tree	:	Pohon makna yang diperoleh dari suatu kalimat.
Parsing	:	Proses analisis kalimat ke dalam beberapa bagian dan menjelaskan peran sintaksis mereka yang lebih kompleks daripada pembacaan sederhana.
Part of Speech	:	Pelabelan kata, frase, hingga klausa sesuai dengan aturan gramatikal. Biasanya juga disingkat menjadi POS.
Precision	:	Metrik untuk mengukur kinerja sistem dalam mendapatkan data yang relevan (disebut juga *positive predicition value*).
PropBank	:	Korpus semantik yang telah memiliki rincian penjelasan proposisi *verb*al dan argumen.
Recall	:	Metrik untuk mengukur kinerja sistem dalam mendapatkan data yang relevan yang terbaca (disebut juga sensitifitas).
Semantik	:	Cabang linguistik yang mempelajari arti/makna yang terkandung pada suatu bahasa, kode, atau representasi lainya.
String	:	Tipe data yang digunakan untuk menyimpan barisan karakter.
XML *Tag*	:	Penanda elemen-elemen, dalam hal ini adalah elemen yang ada di dalam *file* XML.