

DAFTAR ISTILAH

<i>Akurasi</i>	Akurasi merupakan rasio ketepatan sistem dalam mengklasifikasikan dokumen ke kelasnya yang sesuai
<i>Dataset</i>	Kelompok data yang digunakan dalam proses pembangunan dan pengujian <i>classifier</i>
<i>F1-measure</i>	Kombinasi rata-rata dari <i>precision</i> dan <i>recall</i> yang berbanding lurus dengan nilai keduanya
<i>Misspelling</i>	Kesalahan penulisan atau ejaan.
<i>Bayesian network</i>	Metode klasifikasi probabilistik yang mengaplikasikan teorema bayes berdasarkan <i>DAG(Direct Acrylic Graf)</i> .
<i>Part-of-speech tagging</i>	Metode identifikasi jenis kata.
<i>Precision</i>	Rasio jumlah dokumen relevan dengan total jumlah dokumen yang diambil oleh <i>classifier</i> .
<i>Preprocessing</i>	Strategi dan teknik yang saling berkaitan untuk membuat data lebih mudah atau cocok untuk digunakan dalam proses <i>mining</i>
<i>Recall</i>	Rasio dari jumlah dokumen dari koleksi yang ditemukan kembali oleh <i>classifier</i> dengan total jumlah dokumen yang relevan.
<i>Sentiment analysis</i>	Studi yang menganalisis opini seseorang, sentimen, evaluasi bahkan emosi melalui entitas seperti produk, pelayanan, organisasi, individu, topik dan atributnya
<i>DAG(Direct Acrylic Graf)</i>	Sebuah graf yang terdiri dari node dan edge. Pada ujung Edge terdapat mata panah yang menunjukkan arah dari <i>parent</i> atau <i>child</i> dari node.
<i>Stemming</i>	Tahap pengubahan suatu kata menjadi akar atau kata dasarnya dengan cara menghilangkan imbuhan awal atau akhir pada kata tersebut
<i>Lemmatization</i>	Tahap mencari bentuk dasar (lemma) dari kata yang diberikan.
<i>Stop word removal</i>	Tahapan penghapusan kata-kata yang kurang relevan dalam penentuan topik sebuah dokumen
<i>Tokenization</i>	Pemotongan urutan karakter dan sebuah set dokumen yang diberikan menjadi potongan-potongan kata atau karakter yang sesuai dengan kebutuhan sistem