

DAFTAR ISTILAH

No	Istilah	Deskripsi
1	Noise	Noise adalah data yang memiliki nilai yang salah atau nilai yang tidak relevan dengan data yang lain. Noise ini dapat muncul dikarenakan berbagai factor seperti kesalahan pada saat proses pengumpulan data. Noise juga dapat menyebabkan ketidakmaksimalan pada proses pembuatan model machine learning.
2	Outlier	Outlier merupakan data yang memiliki nilai yang jauh dengan data-data yang lain. Nilai outlier biasanya sangat ekstrim perbedaannya dengan data yang lain. Penyebabnya bisa saja kesalahan pada saat pengukuran data. Nilai outlier dapat sangat mempengaruhi dalam proses analisis data.
3	Node	Node merupakan atribut yang menandakan titik penting yang ada pada tree diagram pada algoritma decision tree.
4	TF	TF Merupakan salah satu konsep dalam pembobotan kata yang sangat penting dalam proses analisis teks. TF adalah perhitungan ukuran seberapa sering sebuah kata muncul dalam dokumen. Semakin sering kata muncul maka semakin besar bobotnya dan dianggap penting dan relevan dengan dokumen tersebut.
5	IDF	IDF merupakan salah satu konsep dalam pembobotan kata yang sangat penting dalam proses analisis teks. IDF adalah konsep perhitungan seberapa jarang sebuah kata muncul dalam satu dokumen dengan tujuan untuk menambahkan nilai pada kata yang jarang muncul.