**(TF-IDF)** term frequency–inverse document frequency, is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus

two statistics :a)Term frequency (tf(t,d))- no of times the word repeated in the doc/total number of words in the document

b)Inverse document frequency- helps in determining whether the word is informative and can be considered for extracting conclusions

idf(t,d)=log(total no of documents/no of docs in which word appeared)

**(VSM)**Vector space model or term vector model is an algebraic model for representing text documents (and any objects, in general) as [vectors](https://en.wikipedia.org/wiki/Vector_space) of identifiers, such as, for example, index terms.