Text Mining

Perhitungan Manual untuk TF-IDF

Contoh dokumen

Dokumen 1: "saya suka makan nasi"

* saya: TF = 1/4 = 0.25, IDF = log(3/2) = 0.176, TF-IDF = 0.25 \* 0.176 = 0.044
* suka: TF = 1/4 = 0.25, IDF = log(3/2) = 0.176, TF-IDF = 0.25 \* 0.176 = 0.044
* makan: TF = 1/4 = 0.25, IDF = log(3/1) = 0.477, TF-IDF = 0.25 \* 0.477 = 0.119
* nasi: TF = 1/4 = 0.25, IDF = log(3/3) = 0, TF-IDF = 0.25 \* 0 = 0

Dokumen 2: "saya tidak suka nasi"

* saya: TF = 1/4 = 0.25, IDF = log(3/2) = 0.176, TF-IDF = 0.25 \* 0.176 = 0.044
* tidak: TF = 1/4 = 0.25, IDF = log(3/1) = 0.477, TF-IDF = 0.25 \* 0.477 = 0.119
* suka: TF = 1/4 = 0.25, IDF = log(3/2) = 0.176, TF-IDF = 0.25 \* 0.176 = 0.044
* nasi: TF = 1/4 = 0.25, IDF = log(3/3) = 0, TF-IDF = 0.25 \* 0 = 0

Dokumen 3: "nasi adalah makanan pokok"

* nasi: TF = 1/4 = 0.25, IDF = log(3/3) = 0, TF-IDF = 0.25 \* 0 = 0
* adalah: TF = 1/4 = 0.25, IDF = log(3/1) = 0.477, TF-IDF = 0.25 \* 0.477 = 0.119
* makanan: TF = 1/4 = 0.25, IDF = log(3/1) = 0.477, TF-IDF = 0.25 \* 0.477 = 0.119
* pokok: TF = 1/4 = 0.25, IDF = log(3/1) = 0.477, TF-IDF = 0.25 \* 0.477 = 0.119