Berikut contoh perhitungan manual TF dan IDF pada sebuah paragraph

"Pemerintah mengumumkan kebijakan baru terkait pendidikan. Kebijakan ini diharapkan dapat meningkatkan kualitas pendidikan di seluruh negeri. Selain itu, kebijakan ini juga mencakup pelatihan untuk guru."

Langkah pertama

Pisahkan paragraf menjadi kata-kata.

["pemerintah", "mengumumkan", "kebijakan", "baru", "terkait", "pendidikan", "kebijakan", "ini", "diharapkan", "dapat", "meningkatkan", "kualitas", "pendidikan", "di", "seluruh", "negeri", "selain", "itu", "kebijakan", "ini", "juga", "mencakup", "pelatihan", "untuk", "guru"]

Hitung TF(Term and Frequency)

|  |  |  |
| --- | --- | --- |
| **Kata** | **Frekuensi** | **TF** |
| pemerintah | 1 | 1/25 = 0.04 |
| mengumumkan | 1 | 1/25 = 0.04 |
| kebijakan | 3 | 3/25 = 0.12 |
| baru | 1 | 1/25 = 0.04 |
| terkait | 1 | 1/25 = 0.04 |
| pendidikan | 2 | 2/25 = 0.08 |
| ini | 2 | 2/25 = 0.08 |
| diharapkan | 1 | 1/25 = 0.04 |
| dapat | 1 | 1/25 = 0.04 |
| meningkatkan | 1 | 1/25 = 0.04 |
| kualitas | 1 | 1/25 = 0.04 |
| di | 1 | 1/25 = 0.04 |
| seluruh | 1 | 1/25 = 0.04 |
| negeri | 1 | 1/25 = 0.04 |
| selain | 1 | 1/25 = 0.04 |
| itu | 1 | 1/25 = 0.04 |
| juga | 1 | 1/25 = 0.04 |
| mencakup | 1 | 1/25 = 0.04 |
| pelatihan | 1 | 1/25 = 0.04 |
| untuk | 1 | 1/25 = 0.04 |
| guru | 1 | 1/25 = 0.04 |

Hitung IDF

IDF untuk setiap kata dalam Pargraph ini adalah:

Karena document disini hanya satu paragraph saja maka number of document 1

IDF = log(total documents / number of documents containing the word)

= log(1 / 1)

= log(1)

= 0

IDF(kebijakan) = log(3 / 1) = 0.22

Namun, jika kita memiliki beberapa dokumen, kita akan menghitung IDF sebagai berikut. Misalnya, anggap kita memiliki total 5 dokumen dan kata "kebijakan" muncul dalam 3 dokumen.

Misalkan kita rubah menjadi 3 dokumen

doc1 = "Pemerintah mengumumkan kebijakan baru terkait pendidikan."

doc2 = "Kebijakan ini diharapkan dapat meningkatkan kualitas pendidikan di seluruh negeri."

doc3 = "Selain itu, kebijakan ini juga mencakup pelatihan untuk guru."