

TKI (Temu Kembali Informasi)

IR

Information Retrieval

[BOOLEAN MODEL]

Dokumen diunduh (download) dari <http://temukembaliinformasi.wordpress.com>

Boolean Model

Boolean retrieval merupakan model untuk information retrieval (IR) yang bertujuan untuk membentuk kueri dalam format ekspresi Boolean dari term-term, seperti AND, OR dan NOT.

Seperti namanya Boolean Model, maka setiap term dalam setiap dokumen merupakan vektor 0 (nol) atau 1 (satu). Bernilai 1 jika term tersebut terdapat dalam dokumen, dan bernilai 0 jika term tersebut tidak terdapat dalam dokumen.

Contoh:

Apabila terdapat dokumen:

1. Fahri menghampiri apartemen Maria untuk minta bantuan karena komputer yang digunakan untuk menyusun tesis mengalami error. Maria menuju apartemen tempat tinggal Fahri dan menemukan komputer Fahri kena virus.
2. Selama membantu menyusun tesis, timbul perasaan cinta Maria kepada Fahri.
3.
4.
5.

Kemudian setelah dibentuk tabel berdasarkan frekuensi kemunculan kata/term sebagai berikut:

| Term | dokumen 1 (d1) | d2 | d3 | d4 | d5 |
|----------|----------------|----|----|----|----|
| Fahri | 1 | 1 | 0 | 2 | 0 |
| Maria | 2 | 1 | 3 | 1 | 4 |
| Aisyah | 0 | 0 | 0 | 2 | 3 |
| Noura | 0 | 0 | 2 | 3 | 4 |
| Tesis | 0 | 1 | 2 | 0 | 0 |
| Cinta | 1 | 0 | 0 | 3 | 2 |
| Komputer | 1 | 0 | 2 | 0 | 2 |
| Tesis | 0 | 1 | 2 | 3 | 0 |
| Nikah | 0 | 0 | 1 | 0 | 1 |

Dan query nya adalah Aisyah dan Noura, bukan Fahri. Dokumen manakah yang relevan untuk query tersebut?

Langkah awal, bentuk tabel Boolean dari tabel frekuensi kemunculan kata tersebut menjadi:

Kemudian hitung bentuk vektor Boolean dari kueri.

Query (q) = Aisyah dan Noura, bukan Fahri

= (Aisyah AND Noura) NOT Fahri

= (00011 AND 00111) NOT (11010)

= (00011) AND (00101)

= 00001

Dengan demikian dokumen yang relevan dengan query adalah dokumen 5.