

Tugas Topic Modeling

Dengan menggunakan dataset berita yang telah disediakan, buatlah sebuah model untuk menentukan topik-topik pada dataset.

Manfaatkanlah metode Coherence Model untuk mencari berapa topik yang optimum pada dataset tersebut.

Setelah mengetahui jumlahnya maka cari dan rangkum setiap topik ke dalam satu kalimat.

Jelaskan seperti pada materi,

$(0, '0.086*omicron" + 0.067*varian" + 0.056*negara" + 0.042*covid" + ' '0.023*who" + 0.019*cegah" + 0.016*larang" + 0.016*indonesia" + ' '0.015*afrika" + 0.014*wna"')$

Misalkan indeks 0 karena terdapat, kata who, cegah, larang, indonesia, wna, afrika bisa disimpulkan bahwa Topik A menjelaskan tentang **larangan wna dari afrika** untuk mencegah penyebaran.

Topik 0 - Larangan wna dari afrika

Simpan dalam bentuk pdf.

Masukkan Screenshot hasil kode yang dirasa dibutuhkan pada pdf.

Pertanyaan yang harus dijawab ada 2.

1. Berapa jumlah topik optimum
2. Apa saja topik tersebut dan jelaskan secara singkat kenapa (pada contoh kasus di atas kenapa topik 0 adalah larangan wna dari afrika)