

Analisa Sentimen Netizen Pada Sebuah Berita dan Popularitas Berita dalam Satu Hari

Latar Belakang

Sebuah lembaga riset politik ingin mendapatkan gambaran yang lebih mendalam mengenai narasi berita dan respons publik terhadapnya. Mereka memiliki dua set data:

1. Data Berita: Berisi judul dan isi artikel berita.
2. Data Komentar: Berisi komentar-komentar dan rating bintang yang diberikan oleh pembaca terhadap artikel berita.

Tugas Anda adalah membuat program yang dapat memproses kedua file tersebut dan mengurutkan headline Berita dengan rating bintang tertinggi dan terpopuler berdasarkan kunjungan netizen dalam satu hari ini.

Deskripsi Soal

Fokus dari bagian ini adalah mengidentifikasi topik populer berdasarkan rating comment terhadap berita.

- Pemuatan Data Berita: Baca seluruh data dari ``news_data.csv`` dan ``comment_news.csv``. Simpan data-data tersebut ke dalam dictionary terpisah.
- Pra-pemrosesan Berita: Lakukan langkah-langkah berikut:
 - Cocokkan data dari ``comment_news.csv`` dengan ``news_data.csv`` menggunakan ``idBerita``.
 - Hitung jumlah komentar untuk setiap berita.
 - Hitung rata-rata rating bintang setiap berita ``(rating maksimal 5)``.
- Penyajian Hasil: Tampilkan semua data berita dalam satu tabel (berisikan kolom judul berita, rata-rata rating bintang, dan jumlah komentar) dan urutkan berita dengan rating bintang yang tertinggi.

Struktur Data File

Ada dua file CSV yang akan Anda gunakan:

1. ``news_data.csv``
 - a. Berisi data berita.
 - b. Kolom: ``idBerita``, ``Headline``, ``Content``.
 - c. `idBerita` berfungsi sebagai Primary Key.
2. ``comment_news.csv``
 - a. Berisi data komentar.
 - b. Kolom: ``idKomentar``, ``idBerita``, ``Komentar``, ``Rating``.
 - c. ``idBerita`` di sini berfungsi sebagai Foreign Key, yang menghubungkan setiap komentar ke berita yang relevan.

Catatan

1. File ``stopwords.txt``, ``comment_news.csv``, dan ``news_data.csv`` sudah disediakan, tidak usah membuat lagi.
2. **Kerjakan mandiri dan jangan bertanya kepada teman lainnya**
3. Plagiasi akan di kurangi nilainya
4. Output harus sesuai dengan ketentuan soal
5. Gunakan tipe data ``dictionary`` untuk menyimpan data secara temporary
6. Usahakan menggunakan fungsi untuk penyelesaian soal

Penilaian

1. Kemampuan Membuat Fungsi (35%)
2. Output sesuai dengan ketentuan soal (30%)
3. Menggunakan Streamlit (15%)
4. Kesesuaian UI dengan contoh di soal (20%)

Output/Hasil

Analisis Sentimen & Popularitas Berita				
Menampilkan ID, Headline, Rata-rata Rating, dan Jumlah Komentar, diurutkan dari rating tertinggi.				
	ID Berita	Headline	Rata-rata Rating	Jumlah Komentar
0	108	Sosialisasi Program Bantuan Sosial	4.3300	3
1	101	Rupiah Menguat	4.0000	3
2	103	Pembangunan Infrastruktur di Daerah	3.5000	2
3	105	Pencapaian Ekspor Indonesia	3.5000	2
4	107	Revisi Kebijakan Pajak	3.5000	2
5	110	Pertumbuhan Startup Lokal	3.5000	2
6	104	Keputusan Politik Baru	3.0000	4
7	106	Tantangan Inflasi Global	3.0000	2
8	109	Polemik RUU Kesehatan	2.0000	2
9	102	Kenaikan Harga Bahan Pokok	1.6700	3