

ABSTRAK

Sistem rekomendasi berita personal memainkan peran penting dalam mengatasi kelebihan informasi dengan menyajikan konten berita yang relevan dan terkini kepada pengguna. Namun, sebagian besar studi sebelumnya hanya berfokus pada satu aspek: pendekatan yang semata-mata mengandalkan relevansi semantik sering kali gagal mempertimbangkan kebaruan konten, sementara metode yang hanya bergantung pada aspek kekinian cenderung merekomendasikan artikel berita yang kurang relevan secara kontekstual. Untuk mengatasi keterbatasan ini, kami mengusulkan sebuah kerangka kerja rekomendasi berita inovatif yang mengintegrasikan model transformer RoBERTa-Large dengan fitur tekstual (judul dan abstrak) serta fitur temporal (tanggal publikasi). Pendekatan kami memperkenalkan mekanisme penskoran modular berbasis late fusion, yang secara adaptif menggabungkan skor relevansi dan skor time decay yang diturunkan dari Unix timestamp. Hal ini memungkinkan sistem untuk menyeimbangkan relevansi konten dan kekinian secara fleksibel pada saat inferensi. Untuk meningkatkan keragaman temporal dalam data, kami menggunakan Microsoft News Dataset (MIND), sebuah himpunan data skala besar yang dikumpulkan dari MSN News, yang selanjutnya kami per kaya dengan mengekstraksi tanggal publikasi menggunakan alat web scraping Apify. Hasil eksperimen pada MIND, yang dirata-rata dari 10 kali pengujian, menunjukkan bahwa model RoBERTa-Large yang diusulkan secara konsisten melampaui performa model baseline seperti BERT dan DeBERTaV3, dengan perolehan AUC sebesar 0.7813. Meskipun memiliki tuntutan komputasi yang tinggi, metode kami secara signifikan meningkatkan rekomendasi berita personal dengan secara efektif menangkap relevansi semantik dan dinamika temporal.

Kata kunci: Rekomendasi berita personal, RoBERTa-Large, fitur waktu, time decay, relevansi berita, dataset Microsoft News (MIND)