

Abstraksi

Penelitian ini mengusulkan metode untuk membantu memperbaiki kesalahan ketik, seperti huruf yang salah, huruf yang hilang, dan huruf tambahan. Sampai saat ini, sistem pencarian Lafzi adalah contoh aplikasi yang efektif untuk mencari pertanyaan Arab berdasarkan kesamaan suara. Aplikasi Lafzi membantu memperbaiki kesalahan pengetikan karena bunyi huruf Arab yang hampir sama dengan pelafalan tetapi tidak memperbaiki kesalahan karena pengetikan. Kesalahan ketik dapat mencegah sistem pencarian untuk tidak menampilkan hasil yang diinginkan. Penelitian ini mengusulkan solusi dengan menggunakan metode auto-complete untuk melengkapi trigram yang hilang dan metrik edit distance untuk menghitung nilai diferensiasi antara kueri yang dikoreksi dengan kueri awal. Cara sistem bekerja adalah dengan memisahkan dan menyortir token trigram dari kueri (input pengguna) berdasarkan ayat tertentu. Setiap ayat yang memiliki token trigram yang hilang akan dilengkapi dan ditransformasikan kembali menjadi kueri yang diperbaiki. Setiap kueri yang dikoreksi akan dibandingkan dengan nilai edit distance terhadap kueri awal (input dari pengguna), maka kueri yang dikoreksi akan diambil yang memiliki nilai edit distance terkecil dan akan dibuat sebagai kueri yang disarankan. Evaluasi menunjukkan bahwa metode yang diusulkan menghasilkan nilai recall tertinggi pada 93,40% dan nilai MAP tertinggi pada 86%. Ini mengungguli pendekatan sistem Lafzi sebelumnya yang mencapai nilai recall sebesar 85,23% dan MAP sebesar 79,83%.