

# IMPLEMENTASI ALGORITMA TF-IDF DAN COSINE SIMILARITY DALAM MENENTUKAN KESALAHAN KUTIPAN JURNAL ILMIAH

## ABSTRAK

Sistem pemeriksa kesesuaian kutipan jurnal ilmiah berbasis website menggunakan algoritma TF-IDF dan *Cosine Similarity* untuk mencari tingkat kemiripan teks pada kutipan jurnal ilmiah adalah penelitian yang diteliti karena adanya ketidaksesuaian antara kutipan dan sumber kutipan yang menjadi referensi suatu jurnal ilmiah. Metode pengembangan yang digunakan adalah metode RUP (*Rational Unified Process*). Proses pemeriksaan teks inputan dimulai dari text preprocessing, lalu pembobotan TF-IDF (*Term Frequency Inverse Document Frequency*) dan kemiripan *Cosine Similarity*. Berdasarkan pengujian sistem dengan menggunakan data jurnal ilmiah Bahasa Indonesia online, kemiripan teks menghasilkan tiga kategori, yaitu jika kemiripan teks bernilai 0-0,3 maka teks tersebut tidak similar, jika kemiripan teks bernilai 0,3-0,6 maka teks tersebut mendekati similar, jika kemiripan teks bernilai 0,6-1,0 maka teks tersebut similar. Dari data yang diuji sebanyak 50 inputan teks kutipan dan sumber kutipan maka nilai akurasi teks similar dan mendekati similar yaitu 70% dan tidak similar yaitu 30% karena nilai pembobotan sangat mempengaruhi dalam analisa perhitungan algoritma dalam menentukan tingkat similaritas suatu dokumen.

Kata Kunci: Text Preprocessing, TF-IDF, Cosine Similarity, Jurnal Ilmiah

