

**PERINGKASAN TEKS OTOMATIS PADA JURNAL BERBAHASA  
INDONESIA DENGAN METODE TERM FREQUENCY-INVERSE  
DOCUMENT FREQUENCY (TF-IDF)**

M. Fauzi Adi Saputra<sup>1</sup>, Bagus Setya Rintyarna<sup>2</sup>, Wiwik Suharso<sup>3</sup>

Program Studi Teknik Informatika, Fakultas Teknik

Universitas Muhammadiyah Jember

Email: [mfauzi.ads09@gmail.com](mailto:mfauzi.ads09@gmail.com)

**ABSTRAK**

Perkembangan teknologi internet menyebabkan perkembangan jumlah dokumen digital di internet semakin meningkat. Tetapi dengan keterbatasan manusia untuk memproses informasi baik dari sumber buku, jurnal atau sumber informasi yang lain, maka pembaca memerlukan usaha lebih untuk memahami sebuah teks. Solusi dari masalah tersebut adalah dengan menyediakan atau membuat sebuah ringkasan. Pada penelitian ini peneliti menggunakan automatic summarization. Dan peneliti akan berfokus pada pembuatan ringkasan teks berbasis ekstraktif dengan menggunakan metode Term Frequency-Inverse Document Frequency. TF-IDF adalah sebuah metode pembobotan kata, bobot tersebut yaitu ukuran statistik yang mencerminkan betapa pentingnya sebuah kata bagi sebuah koleksi dokumen atau corpus. Penelitian ini menggunakan 9 jurnal yang berasal dari 3 Program Studi di Universitas Muhammadiyah Jember, dimana setiap jurnal memiliki 5 hasil yang dibedakan berdasarkan Compression Rate (CR) yaitu 50%, 40%, 30%, 20% dan 10%. Peneliti melakukan dua metode evaluasi, yaitu ekstrinsik (pengumpulan pendapat dari responden) dan insterinsik (analisa secara langsung pada ringkasan). Dari hasil evaluasi diperoleh peringkasan terbaik dari kelima hasil terdapat pada ringkasan dengan CR 50% dengan nilai precision, recall, dan f-measure secara berurutan adalah 0,6176, 0,6429, dan 0,6299.

**Kata Kunci:** peringkasan teks ekstraktif, TF-IDF, compression rate.

**AUTOMATIC TEXT SUMMARIZATION IN INDONESIAN LANGUAGE  
JOURNALS USING THE TERM FREQUENCY-INVERSE DOCUMENT  
FREQUENCY (TF-IDF) METHOD**

M. Fauzi Adi Saputra<sup>1</sup>, Bagus Setya Rintyarna<sup>2</sup>, Wiwik Suharso<sup>3</sup>

Informatics Engineering Study Program, Faculty of Engineering

Muhammadiyah University of Jember

Email: [mfauzi.ads09@gmail.com](mailto:mfauzi.ads09@gmail.com)

**ABSTRACT**

The development of internet technology causes the development of the number of digital documents on the internet to increase. But with human limitations to process information either from books, journals or other sources of information, readers need more effort to understand a text. The solution to this problem is to provide or create a summary. In this study, researchers used automatic summarization. And the researcher will focus on making extractive-based text summaries using the Term Frequency-Inverse Document Frequency method. TF-IDF is a word weighting method, the weight is a statistical measure that reflects how important a word is to a collection of documents or corpus. This study uses 9 journals from 3 Study Programs at the University of Muhammadiyah Jember, where each journal has 5 results that are distinguished by Compression Rate (CR), namely 50%, 40%, 30%, 20% and 10%. Researchers conducted two evaluation methods, namely extrinsic (collecting opinions from respondents) and intrinsic (analysis directly on the summary). From the evaluation results, the best summary of the five results is found in the summary with a CR of 50% with precision, recall, and f-measure values respectively 0.6176, 0.6429, and 0.6299.

***Keywords:*** *extractive text summarization, TF-IDF, compression rate.*