

**PENGARUH N-GRAM TERHADAP KLASIFIKASI ABSTRAK
TUGAS AKHIR DI FAKULTAS TEKNIK UNIVERSITAS
MUHAMMADIYAH JEMBER MENGGUNAKAN EKSTRAKSI
DAN SELEKSI FITUR DENGAN METODE
K- NEAREST NEIGHBORS (KNN)**

Errisa Lulut Kustiani¹, Hardian Oktavianto², Guruh Wijaya³

^{1,2,3}*Program Studi Teknik Informatika, Fakultas Teknik, Universitas
Muhammadiyah Jember*

Email: errishakustiani@gmail.com, ²hardian@unmuhjember.ac.id,
³guruh.wijaya@unmuhjember.ac.id

ABSTRAK

Penelitian ini bertujuan untuk menganalisis pengaruh penggunaan *N-gram* terhadap klasifikasi Tugas Akhir mahasiswa di Fakultas Teknik Universitas Muhammadiyah Jember. Dalam penelitian ini, metode ekstraksi dan seleksi fitur menggunakan *N-gram* diimplementasikan bersama dengan algoritma *K-Nearest Neighbors* (KNN) untuk meningkatkan akurasi klasifikasi. Data yang digunakan terdiri dari teks abstrak Tugas Akhir yang diolah menjadi representasi fitur menggunakan teknik *N-gram*. Proses seleksi fitur dilakukan untuk memilih fitur-fitur yang paling signifikan dalam menentukan tingkat akurasi Tugas Akhir. Hasil penelitian menunjukkan bahwa penggunaan *N-gram* sebagai teknik ekstraksi fitur mampu meningkatkan kinerja klasifikasi secara signifikan. Evaluasi performa model dilakukan dengan menggunakan metrik akurasi, presisi, dan *recall*. Hasil penelitian menunjukkan bahwa dengan penambahan *N-gram* dan seleksi fitur kemudian diklasifikasikan dengan KNN, nilai akurasi mencapai 90%, presisi 93%, dan *recall* 90%. Sedangkan jika hanya menggunakan pembobotan TF-IDF *vectorizer* dan diklasifikasikan dengan KNN, nilai akurasi, presisi, dan *recall* masing-masing mencapai 100%. Ini menunjukkan bahwa meskipun *N-gram* dapat meningkatkan kinerja klasifikasi, dalam penelitian ini, TF-IDF *vectorizer* dengan KNN memberikan performa terbaik.

Kata kunci: : *N-gram*, klasifikasi teks, ekstraksi fitur, seleksi fitur, *K-Nearest Neighbors* (KNN)

**INFLUENCE OF N-GRAMS ON CLASSIFICATION OF FINAL PROJECT
ABSTRACTS IN THE FACULTY OF ENGINEERING AT
MUHAMMADIYAH UNIVERSITY OF JEMBER USING FEATURE
EXTRACTION AND SELECTION METHODS WITH K-NEAREST
NEIGHBORS (KNN)**

Errisa Lulut Kustiani¹, Hardian Oktavianto², Guruh Wijaya³

^{1,2,3}*Informatics Engineering Study Program, Faculty of Engineering,
Muhammadiyah University of Jember*

Email: errishakustiani@gmail.com, hardian@unmuhjember.ac.id,
guruh.wijaya@unmuhjember.ac.id

ABSTRACT

This study aims to analyze the impact of using N-gram on the classification of final projects by students at the Faculty of Engineering, Universitas Muhammadiyah Jember. In this research, the feature extraction and selection method using N-gram is implemented together with the K-Nearest Neighbors (KNN) algorithm to improve classification accuracy. The data used consists of the abstract texts of final projects, which are processed into feature representations using N-gram techniques. The feature selection process is carried out to select the most significant features in determining the accuracy of the final project classification. The results show that the use of N-gram as a feature extraction technique can significantly improve classification performance. Model performance evaluation is carried out using accuracy, precision, and recall metrics. The research results show that with the addition of N-gram and feature selection, followed by classification using KNN, the accuracy reaches 90%, precision 93%, and recall 90%. Meanwhile, using only TF-IDF vectorizer weighting and classification with KNN, the accuracy, precision, and recall values each reach 100%. This indicates that although N-gram can improve classification performance, in this study, the TF-IDF vectorizer with KNN provides the best performance.

Keywords: *N-gram, text classification, feature extraction, feature selection, K-Nearest Neighbors (KNN)*