

**ABSTRAK****IMPLEMENTASI *TEXT MINING* PADA LIRIK LAGU UNTUK  
KLASIFIKASI LAGU BERDASARKAN EMOSI MENGGUNAKAN  
*LOGISTIC REGRESSION* BERBASIS *WEBSITE***

Galih Arditiya Yunindar

H1D020091

Musik telah menjadi media penting dalam mengekspresikan emosi dan memperkaya pengalaman sosial manusia. Namun, proses interpretasi emosi pada lirik lagu secara manual sering kali tidak akurat dan membutuhkan waktu, terutama pada lirik yang kompleks atau ambigu. Berbagai algoritma, seperti *K-Nearest Neighbor* (K-NN), *Naive Bayes Classifier*, dan *Support Vector Machine* (SVM), telah diterapkan untuk klasifikasi emosi pada lirik lagu. Penelitian sebelumnya menunjukkan bahwa SVM yang dikombinasikan dengan Particle Swarm Optimization (PSO) mencapai akurasi hingga 90%, sementara K-NN dengan seleksi fitur menghasilkan *f-measure* tertinggi sebesar 66,93%, dan *Naive Bayes* mencapai akurasi hingga 45%. Dalam penelitian ini, algoritma *Logistic Regression* yang didukung oleh metode *Term Frequency-Inverse Document Frequency* (TF-IDF) Hasil evaluasi menunjukkan bahwa model dengan transformasi majas memiliki akurasi yang lebih tinggi (93.52%) dibandingkan dengan model tanpa transformasi majas (92.31%), menunjukkan bahwa majas memberikan kontribusi terhadap kekayaan ekspresi emosional yang dapat dikenali oleh model. Model ini menunjukkan hasil yang kompetitif dan dapat dibandingkan dengan SVM yang menggunakan PSO, sekaligus memberikan kinerja yang lebih baik dibandingkan K-NN dan *Naive Bayes*. Implementasi sistem ini berbasis *website* menggunakan *framework Streamlit*, memungkinkan pengguna memasukkan lirik dan mendapatkan prediksi emosi secara interaktif. Penelitian ini memberikan kontribusi dalam analisis emosi musik, serta menawarkan alternatif yang efisien dan aksesibilitas yang lebih luas untuk klasifikasi emosi dalam lirik lagu.

**Kata Kunci:** Emosi, *Logistic Regression*, *Text Mining*, TF-IDF

**ABSTRACT****IMPLEMENTATION OF TEXT MINING ON SONG LYRICS FOR SONG  
CLASSIFICATION BASED ON EMOTION USING WEBSITE-BASED  
LOGISTIC REGRESSION**

Galih Arditiya Yunindar

H1D020091

*Music has become an essential medium for expressing emotions and enriching human social experiences. However, the manual process of interpreting emotions in song lyrics is often inaccurate and time-consuming, especially with complex or ambiguous lyrics. Various algorithms, such as K-Nearest Neighbor (K-NN), Naive Bayes Classifier, and Support Vector Machine (SVM), have been applied for emotion classification in song lyrics. Previous research has shown that SVM combined with Particle Swarm Optimization (PSO) achieves an accuracy of up to 90%, while K-NN with feature selection produces the highest f-measure of 66.93%, and Naive Bayes achieves an accuracy of up to 45%. In this study, the Logistic Regression algorithm, supported by the Term Frequency-Inverse Document Frequency (TF-IDF) method, is used to improve the accuracy of emotion classification in song lyrics. Evaluation results show that the model with transformation figurative language achieves a higher accuracy (93.52%) compared to the model without transformation figurative language (92.31%), indicating that figurative language not only adds complexity to the text but also contributes to the richness of emotional expression that can be recognized by the model. This model shows competitive results and can be compared to SVM with PSO, while providing better performance than K-NN and Naive Bayes. The system implementation is web-based using the Streamlit framework, allowing users to input lyrics and get interactive emotion predictions. This research contributes to the analysis of music emotions and offers an efficient alternative with wider accessibility for emotion classification in song lyrics.*

**Keywords:** Emotion, Logistic Regression, Text Mining, TF-IDF