

PERBANDINGAN ALGORITMA *RANDOM FOREST*, *SUPPORT VECTOR MACHINE*, DAN *MAXIMUM LIKELIHOOD* PADA KLASIFIKASI *LAND USE LAND COVER* MENGGUNAKAN *REMOTE SENSING DATA*

Oleh

Kadek Prima Giant Marta Dinata, 2015101034

Jurusan Teknik Informatika

ABSTRAK

Penelitian ini bertujuan untuk membandingkan kinerja tiga algoritma klasifikasi, yaitu *Random Forest* (RF), *Support Vector Machine* (SVM), dan *Maximum Likelihood Classification*, dalam konteks klasifikasi *Land Use Land Cover* (LULC) menggunakan data *Remote Sensing Sentinel 2 Level 2A* pada area Chachoengsao, Thailand. Pada penelitian ini digunakan 1500 data dari 5 kelas yang berbeda, dan tiap kelasnya terdiri dari 300 data. Dari 1500 data tersebut dibagi lagi menjadi set pelatihan dan pengujian dengan menggunakan *K-Fold Cross Validation*. Penggunaan data *Remote Sensing Sentinel 2 Level 2A* memastikan analisis yang komprehensif terhadap karakteristik tutupan lahan. Penggunaan *K-Fold Cross Validation* meningkatkan kekokohan model dengan melatih dan menguji pada subset dataset yang beragam. Selanjutnya, model dibuat berdasarkan pembagian data *K-Fold* tersebut. Untuk setiap model yang telah dibuat, akurasi, presisi, *recall*, dan *F1-score* dari tiap model dibandingkan untuk menilai model mana yang paling efektif. Dilakukan pengujian terhadap gambar untuk melihat model mana yang paling akurat. Hasil penelitian menunjukkan bahwa *Maximum Likelihood Classification model* memiliki akurasi tertinggi dibandingkan dengan *Support Vector Machine model* dan *Random Forest model*. Akurasi *Random Forest model* mencapai 95%, sedangkan *Maximum Likelihood Classification model* mencapai 92%, dan *Support Vector Machine model* mencapai 93% pada proses pengujian. Temuan ini mengindikasikan bahwa *Random Forest model* memberikan performa yang lebih baik untuk kasus ini dibandingkan dengan *Support Vector Machine model* dan *Maximum Likelihood Classification model*.

Kata Kunci: *Machine Learning*, *Remote Sensing*, Klasifikasi Penggunaan dan Tutupan Lahan, Sentinel 2, *K-fold Cross Validation*

**COMPARISON OF RANDOM FOREST, SUPPORT VECTOR MACHINE,
AND MAXIMUM LIKELIHOOD ALGORITHMS ON LAND USE LAND
COVER CLASSIFICATION USING REMOTE SENSING DATA**

By

Kadek Prima Giant Marta Dinata, 2015101034

Informatic Engineering Major

ABSTRACT

This research aims to compare the performance of three classification algorithms, namely Random Forest (RF), Support Vector Machine (SVM), and Maximum Likelihood Classification, in the context of Land Use Land Cover (LULC) classification using Remote Sensing Sentinel 2 Level 2A data in the Chachoengsao area, Thailand. The study utilized 1500 data points from 5 different classes, with each class consisting of 300 data points. These 1500 data points were further divided into training and testing sets using K-Fold Cross Validation. The use of Remote Sensing Sentinel 2 Level 2A data ensured a comprehensive analysis of land cover characteristics. The implementation of K-Fold Cross Validation enhanced the robustness of the models by training and testing on diverse subsets of the dataset. Subsequently, models were constructed based on the K-Fold data divisions. For each model, accuracy, precision, recall, and F1-score were compared to evaluate which model was the most effective. Image testing was conducted to determine the most accurate model. The research findings indicated that the Random Forest model outperformed the Support Vector Machine model and the Maximum Likelihood Classification model. The accuracy of the Random Forest model reached 95%, while the Maximum Likelihood Classification model achieved 92%, and the Support Vector Machine model achieved 93% during the testing process. These results suggest that the Random Forest model provided better performance for this case compared to the Support Vector Machine model and the Maximum Likelihood Classification model.

Keywords: *Machine Learning, Remote Sensing, Land Use Land Cover Classification, Sentinel 2, K-fold Cross Validation*