

MACHINE LEARNING-BASED RICE PHENOLOGY MONITORING WITH SATELLITE-DERIVED VEGETATION INDICES

By

Komang Harry Sudana, 2115101053

Informatics Engineering Department

ABSTRACT

Thailand's rice sector faces significant challenges in maintaining productivity, including low yields, labour shortages, and limited water resources. Thailand needs accurate data to guide effective farmland management and address emerging issues as a major global food producer. A comprehensive understanding of crop phenology, especially in rice, is essential to increase production and ensure sustainable management of agricultural resources. This research aims to develop an integrated approach to monitor rice phenology in Chachoengsao Province, Thailand, using multiple vegetation indices, including NDVI, EVI, LSWI, and BSI, combined with machine learning techniques. This study utilized three machine learning models: Support Vector Machine (SVM), Random Forest, and XGBoost. Sentinel-2 satellite imagery with a spatial resolution of 10 meters was used as the dataset, which incorporated 50 position points of rice plants for vegetation index extraction. The analysis showed characteristic patterns in vegetation indices (NDVI, EVI, LSWI, and BSI) corresponding to key phenological stages, where Start of Season (SOS) is characterized by increasing NDVI, EVI, and LSWI with decreasing BSI, Peak of Season (POS) shows peak values except the lowest BSI and End of Season (EOS) features decreasing NDVI, EVI, and LSWI with increasing BSI, reflecting the ageing and drying of rice plants. Among the three models developed, SVM showed the highest performance, achieving 99.4% accuracy, with macro average precision, macro average recall, and macro average f-1 score of 99.3%, 99.4%, and 99.3%, respectively. XGBoost ranked second with 98.8% on all evaluation metrics, while Random Forest ranked last with 97.6% accuracy, 97.9% for macro average precision, macro average recall, and macro average f-1 score. These findings demonstrate the effectiveness of machine learning in phenology analysis and highlight SVM as the most reliable model compared to Random Forest and XGBoost.

Keywords: Machine Learning, Vegetation Index, Crop Phenology, Satellite Imagery

MONITORING FENOLOGI PADI BERBASIS *MACHINE LEARNING* DENGAN INDEKS VEGETASI TURUNAN CITRA SATELIT

Oleh

Komang Harry Sudana, 2115101053

Jurusan Teknik Informatika

ABSTRAK

Sektor beras Thailand menghadapi tantangan yang signifikan dalam mempertahankan produktivitas, termasuk hasil panen yang rendah, kekurangan tenaga kerja, dan sumber daya air yang terbatas. Thailand membutuhkan data yang akurat untuk memandu pengelolaan lahan pertanian yang efektif dan mengatasi masalah-masalah yang muncul sebagai produsen pangan utama dunia. Pemahaman yang komprehensif tentang fenologi tanaman, terutama padi, sangat penting untuk meningkatkan produksi dan memastikan pengelolaan sumber daya pertanian yang berkelanjutan. Penelitian ini bertujuan untuk mengembangkan pendekatan terpadu untuk memantau fenologi padi di Provinsi Chachoengsao, Thailand, dengan menggunakan berbagai indeks vegetasi, termasuk NDVI, EVI, LSWI, dan BSI, yang digabungkan dengan teknik *machine learning*. Penelitian ini menggunakan tiga model *machine learning*: *Support Vector Machine (SVM)*, *Random Forest*, dan *XGBoost*. Citra satelit Sentinel-2 dengan resolusi spasial 10-meter digunakan sebagai dataset, yang menggabungkan 50 titik posisi tanaman padi untuk ekstraksi indeks vegetasi. Hasil analisis menunjukkan pola karakteristik indeks vegetasi (NDVI, EVI, LSWI, dan BSI) yang sesuai dengan tahapan fenologi utama, di mana *Start of Season (SOS)* ditandai dengan peningkatan NDVI, EVI, dan LSWI dengan penurunan BSI, *Peak of Season (POS)* menunjukkan nilai puncak kecuali BSI terendah, dan *End of Season (EOS)* memiliki ciri-ciri penurunan NDVI, EVI, dan LSWI dengan peningkatan BSI yang mencerminkan penuaan dan pengeringan tanaman padi. Di antara tiga model yang dikembangkan, *SVM* menunjukkan kinerja tertinggi, mencapai akurasi 99,4%, dengan *macro average precision*, *macro average recall*, dan *macro average f-1 score* masing-masing 99,3%, 99,4%, dan 99,3%. *XGBoost* menduduki peringkat kedua dengan 98,8% pada semua metrik evaluasi, sementara *Random Forest* menduduki peringkat terakhir dengan akurasi 97,6%, 97,9% untuk *macro average precision*, *macro average recall*, dan *macro average f-1 score*. Temuan ini menunjukkan efektivitas *machine learning* dalam analisis fenologi dan menyoroti *SVM* sebagai model yang paling dapat diandalkan dibandingkan dengan *Random Forest* dan *XGBoost*.

Kata Kunci: *Machine Learning, Vegetation Index, Crop Phenology, Citra Satelit*