

ABSTRAK

Produksi padi merupakan salah satu indikator penting dalam ketahanan pangan nasional, khususnya di wilayah Sumatera yang menjadi salah satu pemroduksi padi di Indonesia. Hasil produksi padi dipengaruhi oleh beberapa variabel. Penelitian ini bertujuan untuk membangun model prediksi produksi padi di pulau Sumatera menggunakan metode *Support Vector Regression* (SVR). SVR bekerja dengan cara mencari fungsi prediksi yang memastikan sebanyak mungkin titik data berada di dalam batas toleransi galat yang disebut ε -tube. Data yang digunakan berasal dari web resmi Badan Pusat Statistik (BPS) Indonesia dan *Kaggle*. Data ini dibagi menjadi dua bagian yaitu data latih dan data uji. Data latih digunakan untuk melatih model prediksi. Dalam penelitian ini dilakukan percobaan dengan membagi model menjadi 4 model yaitu model regresi linear, model SVR dengan parameter default, model dengan pengaturan *hyperparameter* menggunakan *GridSearch*, dan model SVR dengan pengaturan *hyperparameter* menggunakan *RandomizedSearch*. Berdasarkan keempat model tersebut akan dilihat model yang terbaik dalam memprediksi produksi padi. Evaluasi model dilakukan menggunakan metrik MSE dan *R-Squared*, serta analisis kontribusi fitur menggunakan *Permutation Feature Importance* (PFI). Hasil penelitian menunjukkan bahwa model SVR dengan pengaturan *hyperparameter* menggunakan *GridSearch* memberikan performa terbaik dibandingkan model lainnya.

Kata kunci: *Support Vector Regression*, prediksi produksi padi, *hyperparameter tuning*, *Permutation Feature Importance*.

ABSTRACT

Rice production is an important indicator of national food security, particularly in Sumatra, one of Indonesia's rice producers. Rice production is influenced by several variables. This study aims to develop a rice production prediction model on the island of Sumatra using the Support Vector Regression (SVR) method. SVR works by finding a prediction function that ensures as many data points as possible fall within an error tolerance limit called the ε -tube. The data used comes from the official website of the Indonesian Central Statistics Agency (BPS) and Kaggle. This data is divided into two parts: training data and test data. The training data is used to train the prediction model. In this study, an experiment was conducted by dividing the model into four models: a linear regression model, an SVR model with default parameters, a model with hyperparameter settings using GridSearch, and an SVR model with hyperparameter settings using RandomizedSearch. Based on these four models, the best model in predicting rice production will be seen. Model evaluation was carried out using the MSE and R-Squared metrics, as well as feature contribution analysis using Permutation Feature Importance (PFI). The results of the study show that the SVR model with hyperparameter settings using GridSearch provides the best performance compared to other models.

Keywords: Support Vector Regression, rice production prediction, hyperparameter tuning, Permutation Feature Importance