INTISARI

EVALUASI KINERJA *SUPPORT VECTOR MACHINE* (SVM) DALAM PREDIKSI KANKER PARU-PARU

Oleh

BRILLIANT HIBATULLAH EROVAND

E-mail : brillianthibatullahlpg@gmail.com

Penelitian ini bertujuan untuk mengevaluasi kinerja algoritma *Support Vector Machine* (SVM) dalam memprediksi kanker paru-paru menggunakan dataset publik dari Kaggle yang terdiri atas 5.000 data pasien dengan 18 atribut fitur. Metode penelitian menggunakan kerangka kerja CRISP-DM yang meliputi tahapan *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*. Pada tahap *data preparation*, dilakukan transformasi data, normalisasi menggunakan metode *Min-Max*, serta seleksi fitur dengan *Recursive Feature Elimination* (RFE). Pemodelan dilakukan menggunakan algoritma SVM, yaitu dengan kernel *Linear* dan *Radial Basis Function* (RBF). Hasil evaluasi menunjukkan bahwa kernel *Linear* menghasilkan akurasi sebesar 89%, sedangkan kernel *RBF* mencapai 89,6%. Setelah dilakukan penyesuaian *hyperparameter* dengan *GridSearchCV* dan *cross-validation*, model SVM dengan kernel *RBF* mengalami peningkatan akurasi hingga 91,7%, dengan nilai *precision*, *recall*, dan *F1-score* diatas 90% pada kedua kelas target. Penelitian ini membuktikan bahwa SVM, khususnya dengan kernel *RBF* dan parameter yang dioptimalkan, mampu memberikan performa klasifikasi yang unggul serta berpotensi diterapkan pada sistem pendukung keputusan medis untuk deteksi dini kanker paru-paru.

Kata Kunci: *Support Vector Machine*, Prediksi, Kanker Paru-Paru, *Machine Learning*, CRISP-DM

ABSTRACT

PERFORMANCE EVALUATION OF SUPPORT VECTOR MACHINE (SVM) IN LUNG CANCER PREDICTION

By

BRILLIANT HIBATULLAH EROVAND

E-mail : brillianthibatullahlpg@gmail.com

This study aims to evaluate the performance of the Support Vector Machine (SVM) algorithm in predicting lung cancer using a public dataset from Kaggle consisting of 5,000 patient records with 18 feature attributes. The research method applies the CRISP-DM framework, which includes the stages of business understanding, data understanding, data preparation, modeling, evaluation, and deployment. In the data preparation stage, data transformation, normalization using the Min-Max Scaling method, and feature selection through Recursive Feature Elimination (RFE) were performed. The modeling process utilized two SVM kernel types, namely Linear and Radial Basis Function (RBF). Evaluation results indicate that the Linear kernel achieved an accuracy of 89%, while the RBF kernel reached 89.6%. After hyperparameter tuning using GridSearchCV and cross-validation, the SVM model with the RBF kernel achieved an improved accuracy of 91.7%, with precision, recall, and F1-score values exceeding 90% for both target classes. The findings demonstrate that the SVM algorithm, particularly with the optimized RBF kernel, provides superior classification performance compared to previous studies. The resulting model has the potential to serve as a foundation for developing medical decision support systems aimed at early detection of lung cancer efficient, and assist medical professionals in clinical decision-making.

Keywords*: Support Vector Machine, Prediction, Lung Cancer, Machine Learning, CRISP-DM*