

Abstract

Cancer is a disease that occurs due to the uncontrolled growth of abnormal cells causing body tissue damage. This disease is considered as a deadly disease. In 2019, 1700 deaths occur every day due to cancer [1]. Some effective anticancer agents are known to cause temporary to chronic toxic effects. There are several compounds that have the potential to become anticancer drugs, one of them is indenopyrazole. The goal of this research is to implement simulated annealing and support vector machine method in the QSAR study to predict the activity of indenopyrazole derivatives as anticancer drugs. Simulated annealing is used for feature selection and support vector machine is used for model development. In this research, we used three kernel models for SVM, namely SVM with RBF kernel, SVM with linear kernel, and SVM with the polynomial kernel. From three models that were regressed, SVM with RBF kernel has parameter $C = 10$, $\gamma = \text{scale}$ and $\epsilon = 0.1$ produce R^2 score train and test 0.79 and 0.60, respectively. SVM with linear kernel has parameter $C = 1000$, $\text{degree} = 1$ and $\epsilon = 0.1$ produce R^2 score train and test 0.61 and 0.63, respectively. SVM with polynomial kernel has parameter $C = 1000$, $\text{degree} = 2$ and $\epsilon = 0.1$ produce R^2 score train and test 0.72 and 0.50, respectively. Based on the validation results, only model with RBF kernel which parameters satisfy all the criteria. From the result we can conclude that the model with RBF kernel is the best model and acceptable.

Keywords : anticancer, indenopyrazole derivatives, QSAR study, Simulated Annealing, Support Vector Machine