

## ABSTRACT

The improvement in education quality in Indonesia significantly impacts the Human Development Index (HDI). However, the data pattern between education quality and HDI tends to be random, making parametric regression less effective. This study employs a nonparametric kernel regression model with the *Nadaraya-Watson* estimator, optimizing *bandwidth* using *Cross-Validation* (CV) and *Generalized Cross-Validation* (GCV). The *triangle* kernel function was selected due to its efficiency. The results show that the GCV method provides the best model with  $R^2 = 86,32\%$  and MAPE of 1,94%, while the CV method yields the best model with  $R^2 = 84,73\%$  and MAPE of 1,96%. Using GCV, it was found that average years of schooling and expected years of schooling contribute 86,32% to the HDI, with the remaining 13,68% influenced by factors outside the scope of this study. Furthermore, the MAPE value of 1,94% on the test data suggests that the model optimized with GCV performs excellently in forecasting. Overall, the kernel regression model with GCV optimization proves to be more accurate in analyzing the relationship between education quality and HDI, making it a relevant tool for data-driven policy formulation.

**Keywords:** education quality, HDI, multivariable kernel regression, *Cross-Validation*, *Generalized Cross-Validation*, *Nadaraya-Watson*, *triangle* kernel function.