

## INTISARI

### PEMILIHAN MODEL PADA REGRESI RIDGE KERNEL MENGUNAKAN *CROSS VALIDATION*

Oleh

NIKEN LISCA AGGYTA AYUNINGRUM

19/448817/PPA/05900

Regresi ridge kernel merupakan metode yang menggabungkan regresi ridge dengan kernel. Dalam metode kernel terdapat parameter (*bandwidth*),  $\alpha$ , yang berpengaruh pada nilai fungsi kernel. Beberapa fungsi kernel, misalnya Gaussian, Tophat, Linear, Epanechnikov, dan Exponensial. Pemilihan model pada regresi ridge kernel dimulai dengan penentuan model estimator  $\beta$ , pemilihan *bandwidth*,  $\alpha$  dan *shrinkage parameter*,  $\lambda$  menggunakan *cross validation*. Selanjutnya, melakukan simulasi untuk menentukan model yang terbaik. Simulasi yang menghasilkan model dengan RMSE paling kecil dipilih sebagai model optimal. Implementasi pemilihan model pada regresi ridge kernel dilakukan terhadap data iklim harian pengamatan Stasiun Klimatologi Sleman pada tanggal 1 Oktober 2020 sampai dengan 30 Juni 2021, yaitu pada prediksi curah hujan yang dipengaruhi oleh beberapa faktor sebagai prediktor. Berdasarkan hasil analisis data, diperoleh bahwa model regresi ridge kernel optimal untuk prediksi curah hujan dengan RMSE terkecil = 21.18 ketika jenis fungsi kernel yang digunakan Gaussian, nilai  $\alpha = 1.92$ , dan nilai  $\lambda = 5$ .

## **ABSTRACT**

### **MODEL SELECTION IN KERNEL RIDGE REGRESSION USING *CROSS VALIDATION***

By

NIKEN LISCA AGGYTA AYUNINGRUM

19/448817/PPA/05900

Kernel ridge regression is a method that combines ridge regression with the kernel. In the kernel method there is a parameter (bandwidth),  $\alpha$ , which affects the value of the kernel function. Some kernel functions, for example Gaussian, Tophat, Linear, Epanechnikov, and Exponential. The selection of the model in the kernel ridge regression begins with determining the model estimator of  $\beta$ , selecting the bandwidth,  $\alpha$  and shrinkage parameters,  $\lambda$  using cross validation. Next, conducting simulations to determine the best model. The simulation that produces the model with the smallest RMSE is chosen as the optimal model. The implementation of the model selection on the ridge kernel regression was carried out on the daily climatic data observed by the Sleman Climatology Station on October 1, 2020 to June 30, 2021, namely on the prediction of rainfall which was influenced by several factors as predictors. Based on the results of data analysis, it was found that the ridge kernel regression model is optimal for rainfall prediction with the smallest RMSE = 21.18 when the type of kernel function used is Gaussian,  $\alpha$  value = 1.92 and  $\lambda$  value = 5.