

ABSTRACT

THE NONPARAMETRIC REGRESSION OF KERNEL METHOD USING GASSER-MULLER ESTIMATOR ON RAINFALL DATA IN LAMPUNG PROVINCE

By

SHOFIYYAH FAUZIAH SAYUTI

Parametric regression and nonparametric regression are two methods used in regression analysis. In parametric regression, there are several assumptions that must be met, such as the normality of the, the distribution of data that forms a certain pattern, and others, while in nonparametric regression it does not depend on these assumptions. To estimate the nonparametric regression function, smoothing techniques are needed, one of which is the kernel method. This study aims to estimate the nonparametric regression curve using the Gasser-Muller kernel estimator method with the Epanechnikov kernel function using the optimal bandwidth (h) for rainfall data in Lampung Province and to compare the models using Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). From the calculation of the window width formula (bandwidth) from Silverman, the optimal bandwidth (h) value is $hopt_2 = 3.995$ and by using the Generalized Cross Validation (GCV) method, the optimal bandwidth (h) value is $hopt_2 = 5.564$. In addition, by looking at the values of MSE, MAPE, and R^2 , it is found that the bandwidth value (h) $hopt_1 = 3.995$ is better in estimating the curve than $hopt_2 = 5.564$.

Keywords: Regression Analysis, Kernel Method, Gasser-Muller Estimator, Epanechnikov Kernel Function, Bandwidth, Rainfall Data

ABSTRAK

REGRESI NONPARAMETRIK METODE KERNEL DENGAN ESTIMATOR GASSER-MULLER PADA DATA CURAH HUJAN DI PROVINSI LAMPUNG

Oleh

SHOFIYYAH FAUZIAH SAYUTI

Regresi parametrik dan regresi nonparametrik merupakan dua metode yang digunakan dalam analisis regresi. Pada regresi parametrik ada asumsi-asumsi yang perlu dipenuhi seperti kenormalan data, data yang sebarannya membentuk pola tertentu, dan lain sebagainya sedangkan pada regresi nonparametrik tidak bergantung pada asumsi-asumsi tersebut. Untuk mengestimasi fungsi regresi nonparametrik diperlukan teknik pemulusan, salah satunya adalah metode kernel. Penelitian ini bertujuan untuk mendapatkan estimasi kurva regresi nonparametrik dengan metode kernel estimator Gasser-Muller fungsi kernel Epanechnikov menggunakan bandwidth (h) optimal pada data curah hujan di Provinsi Lampung, serta membandingkan model yang telah dibentuk menggunakan Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE), dan Koefisien Determinasi (R^2). Dari hasil perhitungan rumus lebar jedela (bandwidth) dari Silverman diperoleh nilai bandwidth (h) optimal yaitu $h_{opt_1} = 3,995$ dan diperoleh nilai bandwidth (h) optimal yaitu $h_{opt_2} = 5,564$ dengan memakai metode Generalized Cross Validation (GCV). Selain itu dengan melihat nilai MSE, MAPE, dan R^2 , didapat bahwa nilai bandwidth (h) $h_{opt_1} = 3,995$ lebih baik dalam mengestimasi kurva dibandingkan $h_{opt_2} = 5,564$.

Kata kunci: Analisis Regresi, Metode Kernel, Estimator Gasser-Muller, Fungsi Kernel Epanechnikov, Bandwidth, Data Curah Hujan