

ABSTRACT

COMPARISON OF BACKWARD ELIMINATION, FORWARD SELECTION AND STEPWISE REGRESSION METHODS IN THE BEST MULTIPLE LINEAR REGRESSION MODEL SELECTION

By

YULICA DIAH KHOERUNNISA

The best regression model is a model that can explain the behavior of the dependent variable as well as possible, by selecting the independent variables from the many existing independent variables. One of the violations of assumptions in multiple regression analysis is the occurrence of a strong relationship between two independent variables called multicollinearity. The consequence of the multicollinearity problem is that the ordinary small squares estimate still meets the Best Linear Unbiased Estimator (BLUE) requirements, but the estimation becomes inefficient, therefore an improvement is needed to overcome this multicollinearity problem. The purpose of this study is to compare the methods of backward elimination, forward selection and stepwise regression in selecting the best multiple linear regression model in the case of multicollinearity using the coefficient of determination, $R^2_{Adjusted}$ and MSE . The results of this study indicate that the backward elimination, forward selection and stepwise regression methods can overcome multicollinearity and obtain the same linear regression model, so there is no difference in the best regression model in the cases studied in this study.

Keywords : Regression model, Multicollinearity, Backward Elimination, Forward Selection, Stepwise Regression

ABSTRAK

PERBANDINGAN METODE *BACKWARD ELIMINATION, FORWARD SELECTION DAN STEPWISE REGRESSION* DALAM PEMILIHAN MODEL REGRESI LINEAR BERGANDA TERBAIK

Oleh

YULICA DIAH KHOERUNNISA

Model regresi terbaik adalah model yang dapat menjelaskan perilaku variabel tak bebas dengan sebaik-baiknya, dengan memilih variabel-varibel bebas dari sekian banyak variabel bebas yang ada. Salah satu pelanggaran asumsi dalam analisis regresi berganda adalah terjadinya hubungan yang kuat antara dua variabel bebas yang disebut multikolinearitas. Konsekuensi dari masalah multikolinearitas yaitu penaksiran kuadrat kecil biasa masih tetap memenuhi syarat *Best Linear Unbiased Estimator* (BLUE), tetapi penaksiran tersebut menjadi tidak efisien, oleh sebab itu dibutuhkan sebuah perbaikan untuk mengatasi masalah multikolinearitas ini. Tujuan dari penelitian ini adalah membandingkan metode *backward elimination*, *forward selection* dan *stepwise regression* dalam pemilihan model regresi linear berganda terbaik dalam kasus multikolinearitas dengan menggunakan nilai koefisien determinasi, $R^2_{Adjusted}$ dan MSE. Hasil dari penelitian ini menunjukkan bahwa metode *backward elimination*, *forward selection* dan *stepwise regression* dapat mengatasi multikolinearitas dan memperoleh model regresi linear yang sama, sehingga tidak terdapat perbedaan model regresi terbaik pada kasus yang diteliti dalam penelitian ini.

Kata kunci : Model regresi, Multikolinearitas, *Backward Elimination, Forward Selection, Stepwise Regression*