

INTISARI

PENYUSUTAN KOEFISIEN DAN SELEKSI VARIABEL REGRESI MENGUNAKAN METODE MODIFIKASI *BOOTSTRAP LASSO*

Oleh :

Anisa Nur Firdaus

15/378146/PA/16621

Analisis regresi merupakan analisis statistika yang mempelajari hubungan antara satu atau lebih variabel independen dengan variabel dependen. Untuk kasus dimana variabel dependen memiliki dua kemungkinan hasil (*dikotomus*), analisis regresi yang digunakan adalah analisis regresi logistik. Dalam analisis regresi terdapat asumsi yang harus dipenuhi salah satunya adalah tidak ada multikolinearitas pada variabel independen. Apabila asumsi ini tidak terpenuhi menyebabkan model regresi yang terbentuk tidak lagi efisien karena nilai standar error koefisien regresi menjadi sangat besar (*overestimate*).

Pada skripsi ini akan dibahas mengenai Penyusutan Koefisien dan Seleksi Variabel Regresi Menggunakan Metode Modifikasi *Bootstrap LASSO* yang dikembangkan oleh A. Chatterjee dan S.N. Lahiri (2011). Metode Modifikasi *Bootstrap LASSO* merupakan pengembangan dari metode *Bootstrap LASSO* yang sebelumnya dikembangkan dari *LASSO* oleh Tibshirani (1996). Studi kasus ini menggunakan data tipe punggung manusia dan faktor-faktor yang mempengaruhinya. Diperoleh kesimpulan bahwa dengan nilai *threshold* sebesar 0,0053 dan perulangan sebanyak 130 kali, metode Modifikasi *Bootstrap LASSO* lebih efisien untuk mengatasi masalah multikolinearitas dengan memberikan nilai akurasi prediksi model regresi yang lebih baik dibandingkan dengan metode *LASSO* maupun *Bootstrap LASSO*. Hal tersebut dapat dilihat semakin meningkatnya nilai proporsi konkordansi.

Kata kunci : Regresi logistik, multikolinearitas, Modifikasi *Bootstrap LASSO*, *Bootstrap LASSO*, *LASSO*, nilai proporsi konkordansi.

ABSTRACT

**COEFFICIENT SHRINKAGE AND VARIABLE SELECTION
OF REGRESSION USING MODIFIED BOOTSTRAP LASSO METHOD**

By :

Anisa Nur Firdaus

15/378146/PA/16621

Regression analysis is statistical analysis that studies the relationship between one or more independent variables with the dependent variable. For cases where the dependent variable has two possible outcomes (dichotomous), the regression analysis used is logistic regression analysis. In the regression analysis there are assumptions that must be fulfilled one of which is that there is no multicollinearity in the independent variable. If this assumption is not fulfilled, causing the regression model obtained becomes inefficient because the standard error value of the regression coefficient becomes overestimate.

In this paper, will be discussed about Coefficient Shrinkage and Variable Selection Using Modified Bootstrap LASSO method developed by Chatterjee A. and Lahiri S.N. (2011). Modified Bootstrap LASSO method is development of the Bootstrap LASSO method which was previously developed by Tibshirani (1996). This case study uses data on human back type and factors that influence it. It was conclude that using a threshold value of 0,0053 and repetition of 130 times (BM=130), Modified Bootstrap LASSO method is more efficient to overcome multicollinearity problem by giving a better accuration prediction value of regression model than the LASSO and Bootstrap LASSO methods. It can be seen that the concordance proportion value increases.

Keywords : *Logistic regression, multicollinearity, Modified Bootstrap LASSO, Bootstrap LASSO, LASSO, concordance proportion value.*