



INTISARI

REGRESI ADAPTIVE ELASTIC NET UNTUK DATA MULTIKOLINEAR DAN BERDIMENSI TINGGI

Oleh

Fauqi Midad Ni'am

20/459361/PA/20022

Regresi linear merupakan metode statistik yang banyak digunakan untuk memodelkan hubungan antara variabel prediktor dan variabel respon. Namun, pada data dengan karakteristik multikolinearitas dan berdimensi tinggi, metode *Ordinary Least Squares* (OLS) cenderung menghasilkan estimasi koefisien yang tidak stabil dan kurang akurat. Metode regresi terpenalisasi, seperti *Elastic Net* dan *Adaptive LASSO*, dikembangkan untuk mengatasi permasalahan tersebut, tetapi masih memiliki keterbatasan dalam menjaga keseimbangan antara stabilitas estimasi dan seleksi variabel. Penelitian ini bertujuan untuk mengevaluasi kinerja metode *Adaptive Elastic Net* serta membandingkannya dengan *Elastic Net* dan *Adaptive LASSO* dalam membangun model regresi pada berbagai karakteristik data. Analisis dilakukan pada data kasus nyata dan data simulasi yang mencakup data tanpa multikolinearitas dan tidak berdimensi tinggi, data dengan multikolinearitas, serta data berdimensi tinggi (khususnya data *microarray*, $p > n$), termasuk simulasi pembuktian sifat *oracle properties*. Evaluasi kinerja didasarkan pada nilai *Mean Squared Error* (MSE), kemampuan seleksi variabel, nilai *Variance Inflation Factor* (VIF), serta nilai *FI-score* untuk mengukur ketepatan identifikasi koefisien non-nol sejati pada data simulasi. Hasil menunjukkan bahwa *Adaptive Elastic Net* menghasilkan estimasi koefisien yang lebih stabil, seleksi variabel yang lebih efektif, memberikan keseimbangan yang lebih baik antara akurasi prediksi dan kualitas seleksi variabel dibandingkan metode pembanding, serta mampu menangani data dengan multikolinearitas. Selain itu, *Adaptive Elastic Net* menunjukkan performa yang konsisten dalam memenuhi karakteristik *oracle properties*. Dengan demikian, *Adaptive Elastic Net* dapat menjadi metode yang andal dalam analisis regresi terpenalisasi pada data dengan multikolinearitas dan berdimensi tinggi.



ABSTRACT

ADAPTIVE ELASTIC NET REGRESSION FOR MULTICOLLINEAR AND HIGH DIMENSIONAL DATA

By

Fauqi Midad Ni'am

20/459361/PA/20022

Linear regression is a widely used statistical method for modeling the relationship between predictor variables and a response variable. However, in the presence of multicollinearity and high-dimensional data, the Ordinary Least Squares (OLS) method tends to produce unstable and inaccurate coefficient estimates. Penalized regression methods, such as Elastic Net and Adaptive LASSO, have been developed to address these issues, yet they still exhibit limitations in balancing estimation stability and variable selection. This study aims to evaluate the performance of the Adaptive Elastic Net method and to compare it with Elastic Net and Adaptive LASSO in constructing regression models under various data characteristics. The analysis is conducted using real-world data and simulated data, including data without multicollinearity and high dimensionality, data with multicollinearity, and high-dimensional data (particularly microarray data, $p > n$), as well as simulations designed to assess oracle properties. Model performance is evaluated based on the Mean Squared Error (MSE), variable selection capability, the Variance Inflation Factor (VIF), and the F1-score to measure the accuracy of identifying true non-zero coefficients in the simulated data. The results indicate that Adaptive Elastic Net produces more stable coefficient estimates, more effective variable selection, and a better balance between predictive accuracy and selection quality compared to the competing methods, while effectively handling multicollinearity. Furthermore, Adaptive Elastic Net demonstrates consistent performance in satisfying oracle properties. Therefore, Adaptive Elastic Net can be considered a reliable method for penalized regression analysis in the presence of multicollinearity and high-dimensional data.