

**IMPLEMENTASI VARIASI METODE GRADIENT DESCENT
DENGAN PARALLEL COMPUTING DALAM PENGEMBANGAN
R PACKAGE *gradDescent* 4.0**

Oleh

Muhammad Aziz Ashari – muhammad.aziz.ashari@student.upi.edu

1305793

ABSTRAK

Pada penelitian sebelumnya telah dikembangkan *R package gradDescent 3.0* yang berisi algoritma *Gradient Descent* dan 13 variasinya. Setiap algoritma pada *package* tersebut hanya dapat berjalan secara *standalone* sedangkan jumlah data yang diproses di masa mendatang cenderung semakin besar. Fokus dari penelitian ini yaitu mengembangkan *package* sebelumnya dengan memodifikasi 4 dari 13 variasi algoritma yang ada dengan *package* “foreach”, doSNOW, dan “pbMPI” agar memiliki kemampuan dalam memproses data secara *multicore*. Dalam penelitian ini dihasilkan algoritma modifikasi yang kemudian digunakan untuk memprediksi kasus bertipe regresi secara paralel. Untuk menguji *R package* ini dilakukan eksperimen dan simulasi untuk memprediksi *Median House Value*. Didapatkan waktu eksekusi tercepat diraih oleh simulasi “doSNOW” dengan waktu eksekusi 0,01 detik dan nilai galat 80438,66 dengan parameter 4 *core* dan 50 iterasi pada algoritma *AGD* sedangkan nilai RMSE terbaik diraih oleh simulasi yang menggunakan “foreach” dengan nilai galat 70318,16 dengan parameter 4 *core*, 1000 iterasi dengan waktu eksekusi 0,03 detik pada algoritma *AGD*. Berdasarkan hasil penelitian yang dilakukan, *gradDescent 4.0* berhasil dikembangkan. Hasil dari penelitian ini menunjukkan bahwa waktu eksekusi tercepat dan nilai RMSE terbaik di keempat algoritma tersebut diraih oleh mode *parallel computing* dengan skenario *tool*, jumlah *core*, dan iterasi yang berbeda-beda.

Kata kunci: *gradient descent*, regresi linier, *parallel computing*, *R Package*, *California housing*

IMPLEMENTATION OF GRADIENT DESCENT METHOD VARIATION BY USING PARALLEL COMPUTING ON R PACKAGE *gradDescent* 4.0

Arranged by

Muhammad Aziz Ashari – muhammad.aziz.ashari@student.upi.edu

1305793

ABSTRACT

On previous research an R package *gradDescent* 3.0 has been developed which includes Gradient Descent and 13 other variations. Each algorithm on package can only run as a standalone while number of data to process in the future tend to be bigger. Focus of this research is to develop previous package by modifying 4 of 13 variations of algorithm with “foreach”, doSNOW, and “pbMPI” packages to be able to process data at multicore. 4 modified algorithms generated in this research which subsequently used to predict cases of regression type in parallel. The fastest execution time was achieved by a “doSNOW” simulation with a 0.01 second execution time and an error rate of 80438.66 with 4 cores and 50 iterations on the AGD algorithm while the best RMSE score was achieved by a “foreach” simulation with error rate 70318.16 with 4 cores, 1000 iterations with a 0.03 second execution time on the AGD algorithm. Based on result of research, *gradDescent* 4.0 successfully developed. Result of this research shows fastest execution time achieved by parallel computing mode using tool, number of cores, and iterations differently.

Keyword: gradient descent, linear regression, parallel computing, R package, California housing