

***gradDescentR 2.0 : IMPLEMENTASI METODE GRADIENT
DESCENT DAN VARIASINYA DALAM
R PACKAGE***

(STUDI KASUS: FAKTOR KOMPRESIBILITAS GAS)

Dendi Handian (NIM 1200678) dendi@student.upi.edu

ABSTRAK

Machine learning merupakan cabang ilmu komputer yang berfokus pada algoritma yang bisa belajar dari data. Regresi merupakan tugas dari *supervised learning* yang digunakan untuk memprediksi nilai riil berdasarkan variabel *predictor*. Prediksi yang akurat pada regresi ditandai dengan *cost function* yang minimum. Metode *gradient descent* (*GD*) digunakan untuk mencari nilai minimum lokal dari suatu fungsi. Sepanjang waktu *GD* terus berkembang, sehingga banyak dikembangkan berbagai variasi *GD* seperti *Mini-Batch Gradient Descent* (*MBGD*), *Stochastic Gradient Descent* (*SGD*), *Stochastic Average Gradient Descent* (*SAGD*), *Momentum Gradient Descent* (*MGD*), *Accelerated Gradient Descent* (*AGD*), *Adagrad*, *Adadelta*, *RMSprop* dan *Adam*. Penelitian ini berfokus pada pengembangan *R package* bernama *gradDescentR*, yang mengimplementasikan metode *GD* dan variasinya untuk melakukan prediksi pada tugas regresi. *R* merupakan bahasa pemrograman yang populer dalam kasus analisis data, statistik dan *machine learning*. Untuk menguji *R package* ini dilakukan eksperimen dan simulasi untuk mencari atau memprediksi nilai faktor kompresibilitas gas CO₂ berdasarkan parameter tekanan dan suhu yang didapatkan. Penelitian tentang implementasi metode berbasis *GD* untuk memprediksi faktor kompresibilitas gas CO₂ menggunakan *R* sudah pernah dilakukan sebelumnya. Namun pada penelitian tersebut, metode yang diimplementasikan hanya ada empat yaitu *GD*, *MBGD*, *SGD* dan *SAGD*. Penelitian ini akan melanjutkan penelitian tersebut dengan menambahkan metode lain yaitu *MGD*, *AGD*, *Adagrad*, *Adadelta*, *RMSprop* dan *Adam*. Berdasarkan hasil penelitian yang dilakukan, *gradDescentR* berhasil dikembangkan dan sudah terpublikasi pada *Comprehensive R Archive Network* (*CRAN*). Eksperimen dan simulasi *R Package* pada studi kasus faktor kompresibilitas gas CO₂ telah selesai dilakukan dengan hasil rata-rata untuk nilai *RMSE* sebesar 0.008521815 dan waktu eksekusi sebesar 0.14266663 detik.

Kata kunci : *machine learning*, regresi, *gradient descent*, *R*, *R Package*, *CRAN*, faktor kompresibilitas gas.

***gradDescentR 2.0 : IMPLEMENTATION OF GRADIENT
DESCENT METHOD AND ITS VARIANT IN R PACKAGE
(CASE STUDY: GAS COMPRESSIBILITY FACTOR)***

Dendi Handian (NIM 1200678) dendi@student.upi.edu

ABSTRACT

Machine learning is a field of computer science that focus on algorithms that able to learn from data. Regression is one of the supervised learning tasks used to predict a real value (dependent variable) based on predictor value (independent variable). An accurate prediction indicated by a low cost function. Gradient Descent (*GD*) method used to find local minimal of an objective function. Until these days, *GD* had been developed into various algorithms, such as Mini-Batch Gradient Descent (*MBGD*), Stochastic Gradient Descent (*SGD*), Stochastic Average Gradient Descent (*SAGD*), Momentum Gradient Descent (*MGD*), Accelerated Gradient Descent (*AGD*), *Adagrad*, *Adadelta*, *RMSprop* and *Adam*. This research focuses on developing an *R Package* named *gradDescentR*, which it is implementing *GD* method and its variants to perform prediction on regression task. *R* is a popular programming language used for data analysis, statistics, and machine learning. To make use of this *R package*, we propose a case study of CO₂ gas compressibility factor or simply *z-factor*, to predict its value that depends on acquired temperature and pressure. The research about predicting *z-factor* value by using *GD* method in *R* has been done. However, in the research, the implemented method only has four, which is *GD*, *MBGD*, *SGD* and *SAGD*. This research will continue its work by adding more method, such as *MGD*, *AGD*, *Adagrad*, *Adadelta*, *RMSprop* and *Adam*. As the result of this research, *gradDescentR* have been built and published in Comprehensive *R* Archive Network (*CRAN*). The experiment and simulating the *R Package* on CO₂ Gas Compressibility Factor is done with the average result for prediction Root-Mean-Square-Error (*RMSE*) has a value of 0.008521815 and the estimated built time of 0.14266663 seconds.

Keywords : machine learning, regression, gradient descent, *R*, *R package*, *CRAN*, gas compressibility factor.