

ABSTRAK

Nama : Eka Pramudianzah
NIM : 41518010159
Pembimbing TA : Rahmat Budiarto, Dr. Prof
Judul : Komparasi Metode Regresi Linear dan Regresi
Random Forest terhadap Volume Pengangkutan
Sampah

Masalah sampah di Provinsi DKI Jakarta masih menjadi masalah utama yang belum terselesaikan dengan cukup baik karena berbagai faktor seperti jumlah penduduk Jakarta yang terus meningkat setiap tahunnya sehingga tonase sampah juga meningkat serta faktor masyarakat Jakarta yang masih acuh terhadap masalah sampah. Untuk itu, penelitian ini dilakukan dengan menggunakan algoritma *machine learning*. Penelitian ini bertujuan untuk memprediksi pengangkutan sampah di sungai-sungai Jakarta dan menganalisis kinerja algoritma pembelajaran mesin. Algoritma pembelajaran mesin yang digunakan adalah regresi linier dan regresi hutan acak. Prediksi dilakukan dengan menggunakan data volume pengangkutan sampah dari bulan Januari hingga Juni 2021 yang dibedakan berdasarkan penggunaan nilai *outliers* di dalam variabel dependen. Hasil penelitian menunjukkan bahwa bulan Mei merupakan bulan yang memiliki kinerja prediksi yang cukup baik dibandingkan dengan lima bulan lainnya. Namun, berdasarkan performa algoritma maka regresi *random forest* menghasilkan nilai prediksi yang rendah dibandingkan dengan regresi linier karena nilai prediksi dari algoritma regresi *random forest* sebesar 0,82 untuk *mean absolute error* dan 0,81 untuk *mean squared error*.

Kata kunci: Pengangkutan Volume Sampah, *Machine Learning*, Regresi Linear, Regresi *Random Forest*

ABSTRACT

Name : Eka Pramudianzah
Student Number : 41518010159
Counsellor : Rahmat Budiarto, Dr. Prof
Title : Comparison of Linear Regression and Random Forest Regression Methods on Waste Transport Volume

The waste problem in Jakarta Province is still the main problem that has not been resolved quite well due to various factors such as the population, which continues to increase every year so that the tonnage of waste also increases as well as the factor of residents who are still indifferent against waste issues. For this reason, this research was carried out by utilizing two machine learning algorithms, i.e.: Linear Regression and Random Forest Regression. The main aim is to predict the transportation of waste in Jakarta's rivers and analyze the performance of the machine learning algorithms. Experiments are carried out on the waste transport volume data from January to June 2021, with and without considering the outlier values in the dependent variable. Experimental results show that May is the month that has a pretty good predictive performance compared to the other five months. On the algorithm performances, Random Forest Regression provides a low error value compared to Linear Regression, i.e.: 0.82 for the mean absolute error and 0.81 for the mean squared error.

Key words: Waste Transportation, Machine Learning, Linear Regression, Random Forest Regression

UNIVERSITAS
MERCU BUANA