UNIVERSITAS
GADJAH MADA
Aplikasi Machine Learning untuk Memprediksi Kuat Geser Tanah pada Skala Regional
Redo Afdabtur Putra, Rozaqqa Noviandi, S.T., M.Agr., Ph.D.
Universitas Gadjah Mada, 2025 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

Kuat geser tanah, yang ditentukan oleh kohesi (c) dan sudut gesek internal (ϕ), merupakan parameter fundamental dalam geoteknik karena berperan besar terhadap stabilitas lereng, daya dukung fondasi, serta desain struktur penahan tanah. Metode pengujian laboratorium seperti *direct shear*, *triaxial compression*, dan *unconfined compression tests* telah lama digunakan dan diakui akurat, namun memiliki keterbatasan dari sisi waktu, biaya, serta ketersediaan alat dan tenaga ahli. *Machine Learning* (ML) menawarkan pendekatan alternatif berbasis data yang mampu mempelajari hubungan kompleks antara parameter indeks tanah dengan kuat geser. Meskipun demikian, sebagian besar penelitian terdahulu hanya menggunakan dataset tunggal tanpa mempertimbangkan faktor spasial, sehingga kemampuan generalisasi model pada skala regional masih belum banyak dikaji.

Penelitian ini menggunakan data sekunder hasil uji penyelidikan tanah dari Laboratorium Mekanika Tanah DTSL FT UGM, yang dikumpulkan dari 31 laporan soil investigation di Provinsi Jawa Tengah dan Daerah Istimewa Yogyakarta dengan total 31 titik sampel. Dataset tersebut terdiri atas 11 data tanah kohesif dan 20 data tanah non-kohesif, dengan kondisi geologi meliputi *basalt*, *alluvium*, *diorite rock*, hingga *river terrace,* kemudian untuk kondisi litologinya meliputi *Extrusive Intermediate Polimik, Sediment Classic Alluvium, Extrusive Intermediate lava*, dan *Sedimen Limestone*, dan dengan elevasi yang berbeda-beda. Variabel yang digunakan dalam pemodelan meliputi batas Atterberg, kadar air, *void ratio*, *specific gravity*, dan persen butiran halus. Empat algoritma ML dibandingkan dalam penelitian ini, yaitu *Artificial Neural Network* (ANN), *Support Vector Machine* (SVM), *Gaussian Process Regression* (GPR) dan *Extreme Gradient Boosting* (XGBoost), dengan dua skenario: (i) model tanpa input lokasi dan (ii) model dengan input lokasi. Evaluasi kinerja dilakukan menggunakan koefisien determinasi (R^2), *Root Mean Square Error* (RMSE), dan *Mean Absolute Error* (MAE).

Hasil penelitian menunjukkan bahwa korelasi antara variabel prediktor tunggal dan parameter kuat geser tanah (c dan ϕ) relatif rendah untuk sebagian besar algoritma, meskipun XGboost menghasilkan nilai yang cukup (R^2 = 0,227 – 0,805) untuk sudut gesek internal, (R^2 = 0,002 – 0,218) untuk kohesi. Variasi hasil nilai prediksi tersebut disebabkan perbedaan karakteristik data setiap parameter, serta kemampuan masing-masing algoritma menangkap hubungan nonlinear dengan kuat geser tanah, oleh karena itu prediksi dengan kombinasi variabel kelompok perlu dilakukan untuk hasil yang lebih baik. Kombinasi variabel kelompok dan penambahan faktor lokasi mampu meningkatkan performa model secara signifikan. XGBoost memberikan hasil terbaik, dengan nilai R^2 = 0,931, RMSE = 3,81, dan MAE = 2,29 untuk sudut gesek internal (ϕ), dan R^2 = 0,879, RSME = 0,07, dan MAE = 0,04 untuk kohesi (c) pada tanah kohesif. Pada tanah non-kohesif, meskipun akurasi lebih rendah, penambahan parameter lokasi tetap meningkatkan prediksi dengan hasil terbaik R^2 = 0,501, RMSE = 2,55 dan MAE = 1,74 untuk ϕ, dan R^2 = 0,793, RSME = 0,07, dan MAE = 0,02 untuk c. Temuan ini menunjukkan bahwa integrasi variabel lokasi dapat meningkatkan prediksi ML pada skala regional. Kondisi geologi, iklim, dan hidrologi merupakan faktor utama dalam proses pembentukan tanah yang mengontrol sifat fisik maupun mekanik tanah. Variasi pada faktor-faktor tersebut akan menentukan tekstur, struktur, kadar air, serta tingkat pelapukan tanah, yang pada akhirnya berpengaruh langsung terhadap parameter kuat geser tanah, baik kohesi maupun sudut gesek internal tanah.

Kata kunci: Kuat Geser Tanah, Parameter indeks, Tanah Kohesif, Tanah Non-kohesif, *Machine Learning*

UNIVERSITAS
GADJAH MADA

Aplikasi Machine Learning untuk Memprediksi Kuat Geser Tanah pada Skala Regional
Redo Afdabtur Putra, Rozaqqa Noviandi, S.T., M.Agr., Ph.D.
Universitas Gadjah Mada, 2025 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

Soil shear strength, defined by cohesion (c) and internal friction angle (ϕ), is a fundamental parameter in geotechnical engineering as it governs slope stability, bearing capacity, and the design of earth-retaining structures. Conventional laboratory tests such as direct shear, triaxial compression, and unconfined compression have long been recognized for their accuracy; however, these methods are often limited by time, cost, equipment availability, and the need for skilled operators. Machine Learning (ML) offers a data-driven alternative capable of capturing complex relationships between soil index parameters and shear strength. Nevertheless, most previous studies relied on single datasets without incorporating spatial factors, leaving the generalization capability of models at the regional scale largely unexplored.

This study utilizes secondary soil investigation data obtained from the Soil Mechanics Laboratory, Department of Civil and Environmental Engineering, Universitas Gadjah Mada (DTSL FT UGM), comprising 31 investigation reports from Central Java and the Special Region of Yogyakarta, totaling 31 sampling points. The dataset consists of 11 cohesive and 20 non-cohesive soils derived from various geological settings, including basalt, alluvium, diorite rock, and river terrace formations, as well as lithologies such as extrusive intermediate polymic, clastic alluvium, extrusive intermediate lava, and sedimentary limestone, across varying elevations. The input variables include Atterberg limits, water content, void ratio, specific gravity, and percentage of fine particles. Four ML algorithms—Artificial Neural Network (ANN), Support Vector Machine (SVM), Gaussian Process Regression (GPR), and Extreme Gradient Boosting (XGBoost)—were evaluated under two scenarios: (i) models without spatial input and (ii) models incorporating spatial input. Model performance was assessed using the coefficient of determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE).

The results indicate that single-variable correlations with shear strength parameters (c and ϕ) were generally low across most algorithms, although XGBoost achieved moderate accuracy (R^2 = 0.227–0.805 for ϕ and R^2 = 0.002–0.218 for c). These variations arise from differences in the statistical characteristics of each variable and the algorithms' ability to capture nonlinear relationships. Consequently, multi-variable combinations were explored to improve model performance. Incorporating grouped variables and spatial factors significantly enhanced predictive accuracy. XGBoost achieved the best performance for cohesive soils, with R^2 = 0.931, RMSE = 3.81, and MAE = 2.29 for ϕ, and R^2 = 0.879, RMSE = 0.07, and MAE = 0.04 for c. For non-cohesive soils, although the overall accuracy was lower, adding location parameters still improved the predictions, yielding R^2 = 0.501, RMSE = 2.55, and MAE = 1.74 for ϕ, and R^2 = 0.793, RMSE = 0.07, and MAE = 0.02 for c. These findings demonstrate that integrating spatial variables can enhance the regional-scale prediction capability of ML models. Geological, climatic, and hydrological conditions are major factors influencing soil formation processes, which in turn govern the physical and mechanical properties of soils. Variations in these factors determine soil texture, structure, moisture content, and degree of weathering, ultimately affecting both cohesion and internal friction angle.

Keywords: *Soil Shear Strength, Index Parameters, Plastic Soil, Non-Plastic Soil, Machine Learning*