

INTISARI

Air tanah merupakan salah satu sumber utama dalam pemenuhan kebutuhan air bersih, dimana 67% penduduk Kabupaten Trenggalek bergantung pada air tanah. Namun demikian, Kabupaten Trenggalek mengalami kekeringan setiap tahunnya, sehingga dapat mengurangi kuantitas air tanah. Oleh karena itu, informasi potensi air tanah memiliki peran penting dalam manajemen sumber daya air bersih berkelanjutan. Sistem Informasi Geografis (SIG) dapat digunakan untuk pemetaan potensi air tanah dengan memerlukan teknik dengan tingkat objektivitas tinggi. Salah satu teknik yang dapat digunakan adalah algoritma berbasis *artificial intelligence* (AI). Oleh karena itu, penelitian ini bertujuan untuk memetakan potensi air tanah dengan teknik SIG dan pendekatan algoritma berbasis AI.

Penelitian ini menggunakan dua algoritma *machine learning* (ML) dan dua algoritma *deep learning* (DL). Algoritma ML yang digunakan adalah *random forest* (RF) dan *gradient boosting decision tree* (GBDT). Sementara itu, algoritma DL yang digunakan adalah *recurrent neural network* (RNN) dan *convolutional neural network* (CNN). Selain menggunakan algoritma tunggal, penelitian ini juga mengimplementasikan pengintegrasian algoritma dengan skema ML, DL, dan ML-DL menggunakan teknik *stacking learning*. *Meta learning* yang digunakan adalah *logistic regression* (LR). Pemodelan potensi air tanah pada penelitian ini menggunakan 18 parameter yaitu *topographic wetness index* (TWI), kerapatan sungai, *stream power index* (SPI), curah hujan, evapotranspirasi, suhu permukaan tanah, tekanan atmosfer permukaan, tutupan lahan, *normalized difference vegetation index* (NDVI), elevasi, kemiringan lereng, arah lereng, *plan-curvature*, kerapatan pelurusan, jenis geologi, jenis tanah, *complete bouguer anomaly* (CBA), dan gelombang geser rata-rata pada kedalaman 30m (Vs30). Secara garis besar, penelitian ini terdiri dari dua tahapan utama yaitu persiapan data parameter dan pemodelan potensi air tanah. Data dari berbagai sumber dipersiapkan untuk menjadi parameter pemodelan berformat raster dengan resolusi spasial 30m. Kemudian dilakukan proses pemodelan dengan menggunakan algoritma RF, GBDT, RNN, CNN, RF-GBDT, RNN-CNN, dan RF-GBDT-RNN-CNN. Data sebaran mata air sebagai data *training* pada label potensial, dan *random point* pada area jenis geologi dengan porositas rendah sebagai *training* label non-potensial. Delapan metrik digunakan untuk mengevaluasi antar model yaitu *accuracy* (Acc), *sensitivity* (Sn), *specificity* (Sp), *geometric mean* (GM), *balanced accuracy* (BA), *mattew's coefficient correlation* (MCC), dan *kappa accuracy* (KA).

Berdasarkan hasil dari 7 model prediksi, wilayah studi didominasi oleh potensi rendah dengan persentase luasan 22,32% hingga 31,29%. Area dengan potensi tinggi cenderung berada di bagian selatan wilayah studi. Sementara itu area dengan potensi rendah cenderung berada di bagian utara wilayah studi. Model SL-ML merupakan model terbaik pada penelitian ini dengan menghasilkan nilai Acc, Sn, Sp, GM, BA, MCC, dan KA masing-masing adalah 0,957; 0,972; 0,944; 0,958; 0,957; 0,958; 0,915; dan 0,915. Parameter jenis geologi merupakan parameter paling signifikan mempengaruhi semua model prediksi dengan nilai *permutation feature importance* (PFI) sebesar 0,242 hingga 0,349. Hasil penelitian ini dapat menjadi acuan pengembangan penelitian selanjutnya terkait pemodelan potensi air tanah.

Kata kunci : Potensi air tanah, *deep learning*, *machine learning*, dan *stacking learning*

ABSTRACT

Groundwater is one of the main sources for supplying clean water, which 67% of the population of Trenggalek Regency depends on. However, Trenggalek Regency experiences drought annually, which can reduce the quantity of groundwater. Therefore, information on groundwater potential plays a crucial role in the sustainable management of clean water resources. Geographic Information Systems (GIS) can be used for mapping groundwater potential, which requires techniques with a high degree of objectivity. The AI-based algorithm is one such technique that can be employed. This study aims to map groundwater potential using GIS techniques and an AI-based algorithm approach.

This study utilizes two machine learning (ML) algorithms and two deep learning (DL) algorithms. The ML algorithms utilized are random forest (RF) and gradient boosting decision tree (GBDT). Meanwhile, the DL algorithms utilized are recurrent neural network (RNN) and convolutional neural network (CNN). In addition, this study implemented the integration of ML, DL, and ML-DL algorithms using the stacking learning technique. The meta learning used is logistic regression (LR). Groundwater potential modelling in this study uses 18 parameters, namely topographic wetness index (TWI), river density, stream power index (SPI), rainfall, evapotranspiration, soil surface temperature, surface atmospheric pressure, land cover, normalized difference vegetation index (NDVI), elevation, slope gradient, slope direction, plan-curvature, straightness density, geological type, soil type, complete Bouguer anomaly (CBA), and average shear wave velocity at a depth of 30m (Vs30). In general, this study consists of two main stages: parameter data preparation and groundwater potential modelling. Data from various sources were prepared as raster-formatted modelling parameters with a spatial resolution of 30m. Then, modelling was performed using the RF, GBDT, RNN, CNN, RF-GBDT, RNN-CNN, and RF-GBDT-RNN-CNN algorithms. Spring distribution data was used as training data for potential labels, and random points in areas with low porosity geology were used as training data for non-potential labels. Eight metrics were used to evaluate the models: accuracy (Acc), sensitivity (Sn), specificity (Sp), geometric mean (GM), balanced accuracy (BA), Matthew's coefficient correlation (MCC), and kappa accuracy (KA).

According to the 7 model results, the study area is dominated by low potential, with an area percentage ranging from 22.32% to 31.29%. Areas with high potential tend to be located in the southern part of the study area. Meanwhile, areas with low potential tend to be located in the northern part of the study area. The RF-GBDT model is the best model in this study, producing Acc, Sn, Sp, GM, BA, MCC, and KA values of 0.957, 0.972, 0.944, 0.958, 0.957, 0.958, 0.958, 0.915, and 0.915, respectively. Geological type parameters are the most significant parameters influencing all prediction models, with permutation feature importance (PFI) values ranging from 0.242 to 0.349. The results of this study can serve as a reference for further research on groundwater potential modelling.

Keywords: Groundwater potential map, deep learning, machine learning, dan stacking learning