

**METODE NON-PARAMETRIK CLASSIFICATION TREE ANALYSIS (CTA)  
DENGAN TEKNIK DATA MINING UNTUK KLASIFIKASI  
PENGGUNAAN LAHAN MENGGUNAKAN CITRA LANDSAT-8 OLI**

**Cahya Budi Perwitagama**

**12/337168/PGE/00954**

**INTISARI**

Kajian mengenai metode non-parametrik *Classification Tree Analysis* (CTA) menggunakan teknik *data mining* untuk aplikasi penginderaan jauh masih belum banyak dilakukan. Sehingga diperlukan penelitian mengenai kemampuan CTA dalam menangani data yang cukup banyak, dengan memanfaatkan kelebihan CTA untuk aplikasi penginderaan jauh. Kombinasi parameter CTA dan data masukan, serta penerapannya pada dua skema klasifikasi yang berbeda tingkat kerinciannya, memerlukan pengujian terkait dengan tingkat akurasi yang dihasilkan. Penelitian ini bertujuan melakukan simulasi dari beberapa kombinasi parameter untuk mengetahui tingkat akurasi hasil klasifikasi, dan memperoleh pohon keputusan dari hasil KDD. Serta menganalisis akurasi metode non-parametrik CTA dengan teknik *data mining* untuk klasifikasi penggunaan lahan menggunakan citra Landsat-8 OLI, dan menerapkan hasil KDD pada daerah lain.

Klasifikasi diperoleh dengan melakukan simulasi beberapa parameter CTA dan data masukan. Parameter aturan pemisah (*splitting rules*) dalam CTA, yaitu *Ratio*, *Entropy*, dan *Gini*. Parameter pemangkas (*pruning*), yaitu 0%, 1%, 5%, dan 10%. Beberapa masukan data klasifikasi, antara lain adalah citra Landsat-8 tujuh saluran, transformasi citra (NDVI, NDWI, BI, NDBI, dan PCA), serta filter tekstur *variance* dan *mean* (jendela bergerak 3x3 dan 5x5). Data non-spektral, yaitu data ketinggian dan data kemiringan lereng. Dua tingkat skema klasifikasi penggunaan lahan, yaitu Level I (5 kelas) dan Level II (8 kelas). Pohon keputusan yang diperoleh dari hasil pembelajaran dengan akurasi terbaik kemudian diterapkan pada daerah lain yang memiliki karakteristik mirip.

Hasil penelitian menunjukkan bahwa kombinasi parameter terbaik adalah atribut pemisah *Gini*, *pruning* 1%, filter tekstur dengan jendela bergerak 5x5, dan skema Level I yaitu dengan akurasi keseluruhan 96,71%, *kappa* 0,9504, dan waktu proses 3,388 detik. Penerapannya pada daerah lain, menghasilkan akurasi keseluruhan 93,27% dengan *kappa* 0,8923. Tingkat akurasi terbaik yang diperoleh pada daerah penelitian maupun penerapannya pada daerah lain, lebih besar dari 90%. Sehingga diharapkan metode ini dapat menjadi alternatif metode untuk terapan kebijakan penggunaan lahan, dan klasifikasi penggunaan lahan berbasis penutup lahan setara dengan skala 1:100.000.

**Kata Kunci:** *Classification Tree Analysis* (CTA), *data mining*, klasifikasi penggunaan lahan, Landsat-8

**NON-PARAMETRIC METHODS CLASSIFICATION TREE ANALYSIS (CTA)  
WITH DATA MINING TECHNIQUES FOR LAND USE CLASSIFICATION  
USING LANDSAT-8 OLI IMAGERY**

**Cahya Budi Perwitagama**

**12/337168/PGE/00954**

**ABSTRACT**

*The study of non-parametric methods Classification Tree Analysis (CTA) using data mining techniques for remote sensing applications is still much to do, so it require studies on the CTA's ability to handle quite a lot of data, by utilizing CTA advantages for remote sensing applications. The combination of parameter CTA and data input, as well as its application on two different detail levels of land use classification scheme, require testing related to the level of accuracy that is generated. This studies aimed to simulate multiple combinations of parameters to determine the level of accuracy of the classification results, and obtain a decision tree from KDD results. And to analyze the accuracy of non-parametric methods CTA with data mining techniques for land use classification using Landsat-8 OLI, and apply the results of KDD on another area.*

*The classification is obtained by performing simulations on several parameters CTA and data input. There are three splitting rules parameter in CTA, i.e. Ratio, Entropy, and Gini. Pruning parameter, i.e. 0%, 1%, 5%, and 10%. There are several inputs of the classification data, namely seven bands of Landsat-8 OLI imagery, image transformation (NDVI, NDWI, BI, NDBI, and PCA), as well as texture filter variance and texture filter mean (moving window 3x3 and 5x5). Non-spectral data, i.e. elevation data and slope data. And two-level land use classification scheme, i.e. Level I (5 classes) and Level II (8 classes). The decision tree that obtained from the best accuracy of the learning outcomes then applied to another area with similar characteristics.*

*The results showed that the best parameter combination are a splitting attribute Gini, pruning 1%, texture filter with 5x5 moving window, and Level I scheme, with an accuracy of 96.71%, kappa 0.9504, and processing time 3.388 sec. Its application on another area, resulting an overall accuracy 93.27% with kappa 0.8923. The best accuracy rate obtained in this study and its application on another area was greater than 90%. Therefore this method is expected could be an alternative method for land use policy application, and land use classification based on land cover commensurate to a scale of 1:100,000.*

**Keywords:** *Classification Tree Analysis (CTA), data mining, land use classification, Landsat-8*