

EXPLORING THE DECISION TREE METHOD FOR MODELLING URBAN LAND USE CHANGE

Mileva Samardžić-Petrović¹, Suzana Dragičević¹,
Branislav Bajat² and Miloš Kovačević³

¹Spatial Analysis and Modeling Laboratory, Department of Geography, Simon Fraser University, Burnaby, BC, Canada

²Department of Geodesy and Geoinformatics, Faculty of Civil Engineering, University of Belgrade, Belgrade, Republic of Serbia

³Department of Construction Project Management, Faculty of Civil Engineering, University of Belgrade, Belgrade, Republic of Serbia

Land use changes play an important role in interactions between human and physical systems, and have significant impacts on the environment at local, regional and global scales. Land use change is a complex process and so developing dynamic models to represent the process is a challenging task. Decision Trees (DT) is a Machine Learning (ML) method with the capability to extract trends and generate a representative model using historical geospatial data. While DT is used in remote sensing as an image classification method, it is not sufficiently examined in land use science. The main objective of this research study is to examine the capability of DT method to model urban land use change. Various numbers of attributes for three municipalities in the City of Belgrade, Republic of Serbia were used. Land use is represented with nine land use classes for three different time instances for the years 2003, 2007 and 2011. The kappa statistics and weighted Area Under Receiver Operating Characteristic Curve (AUC) were used to compare the model outputs with real land use datasets for year 2011. The maximum obtained values for kappa and weighted AUC indicate that DT is a useful method for modelling urban land use change. Furthermore, the derived classification tree generates information about the relationship between the considered causal factors and land use changes and allows for better understanding of the change process.

Les changements dans l'affectation des terres jouent un rôle important dans les interactions entre les humains et les systèmes physiques et ont des impacts importants sur l'environnement aux échelles locale, régionale et mondiale. Le changement d'affectation des terres est un processus complexe. C'est pourquoi le développement de modèles dynamiques pour représenter ce processus constitue une tâche remplie de défis. L'arbre de décision (AD) est une méthode d'apprentissage machine qui a la capacité d'extraire des tendances et de produire un modèle représentatif qui utilise les données géospatiales historiques. Même si l'arbre de décision est utilisé en télédétection comme méthode de classification, il n'est pas suffisamment examiné dans la science de l'aménagement du territoire. L'objectif principal du présent article de recherche est d'examiner la capacité de la méthode de l'arbre de décision pour modéliser le changement d'affectation des terres urbaines. Nous avons utilisé un nombre varié d'attributs pour trois municipalités dans la ville de Belgrade en République de Serbie. L'utilisation des terres est représentée à l'aide de neuf classes d'utilisation des terres sur trois différentes périodes de temps au cours des années 2003, 2007 et 2011. Les statistiques « kappa » et la courbe pondérée de la fonction d'efficacité du récepteur (CFER) ont été utilisées pour comparer les extraits du modèle à des jeux de données réelles d'utilisation des terres pour l'année 2011. Les valeurs maximales obtenues pour les statistiques kappa et la CFER indiquent que l'arbre de décision est une méthode utile pour modéliser le changement d'affectation des terres urbaines. En outre, l'arbre de classification qui en découle produit de l'information sur la relation entre les facteurs de causalité pris en compte et les changements d'affectation des terres et permet de mieux comprendre le processus de changement.

1. Introduction

The recent expansions of spatial data acquisition technologies, higher computation capacities and increased availability of public and free data have a great impact on the development of data mining strategies. One of the techniques which present the

cornerstone of data mining is machine learning (ML) [Fayyad et al. 1996]. Data mining involves “the extraction of implicit, previously unknown, and potentially useful information from data” [Witten et al. 2011]. Data mining strategies have



Mileva Samardžić-Petrović



Suzana Dragičević
suzanad@sfu.ca



Branislav Bajat



Miloš Kovačević