Environmental Modelling & Software 84 (2016) 240–250

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India)

Binh Thai Pham [a, b, *], Biswajeet Pradhan [c], Dieu Tien Bui [d], Indra Prakash [e], M.B. Dholakia [f]

[a] Department of Civil Engineering, Gujarat Technological University, Nr.Visat Three Roads, Visat, Gandhinagar Highway, Chandkheda, Ahmedabad, 382424, Gujarat, India
[b] Department of Geotechnical Engineering, University of Transport Technology, 54 Trieu Khuc, Thanh Xuan, Ha Noi, Viet Nam
[c] Faculty of Engineering, Geospatial Information Science Research Centre (GIS RC), Department of Civil Engineering, University Putra Malaysia, 43400, Serdang, Selangor Darul Ehsan, Malaysia
[d] Geographic Information System Group, Department of Business Administration and Computer Science, Telemark University College, Hallvard Eikas Plass 1, N-3800, Bø i Telemark, Norway
[e] Department of Science & Technology, Bhaskarcharya Institute for Space Applications and Geo-Informatics (BISAG), Government of Gujarat, Gandhinagar, India
[f] Department of Civil Engineering, LDCE, Gujarat Technological University, Ahmedabad, 380015, Gujarat, India

ARTICLE INFO

Article history:
Received 20 March 2016
Received in revised form
23 May 2016
Accepted 2 July 2016

Keywords:
Landslides susceptibility assessment
Machine learning
Uttarakhand
India

ABSTRACT

Landslide susceptibility assessment of Uttarakhand area of India has been done by applying five machine learning methods namely Support Vector Machines (SVM), Logistic Regression (LR), Fisher's Linear Discriminant Analysis (FLDA), Bayesian Network (BN), and Naïve Bayes (NB). Performance of these methods has been evaluated using the ROC curve and statistical index based methods. Analysis and comparison of the results show that all five landslide models performed well for landslide susceptibility assessment (AUC = 0.910–0.950). However, it has been observed that the SVM model (AUC = 0.950) has the best performance in comparison to other landslide models, followed by the LR model (AUC = 0.922), the FLDA model (AUC = 0.921), the BN model (AUC = 0.915), and the NB model (AUC = 0.910), respectively.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Landslide is one of the most serious geo-hazards causing the loss of life and property all over the world, therefore, landslide susceptibility assessment and mitigation of its harmful impactions has been turning into urgent tasks to government and non-government agencies (Althuwaynee et al., 2012). Landslide susceptibility map is known as a useful tool for landslide hazard management through land use planning and better decision making in landslide prone areas (Akgun, 2012). Machine learning approaches are considered more efficient than other approaches such as expert's opinion based methods and analytic methods for spatial prediction of landslides (Pradhan, 2013). Main principle of these approaches is that landslide susceptibility is assessed using machine learning algorithms to analyze the spatial relationship between past landslide events and a set of conditioning factors from which the potential probability of landslide occurrence is determined (Chen et al., 2015; Guzzetti, 2006).

Many types of machine learning algorithms have been developed and applied for producing landslide susceptibility maps in many regions of the world. Zare et al. (2013), Pradhan and Lee (2010b), and Conforti et al. (2014) utilized artificial neural networks which are based on the biological neural networks to predict spatially landslide distributions. Whereas, Tien Bui et al. (2012b), Lee et al. (2013), and Jebur et al. (2015) applied evidential belief functions to generate landslide susceptibility maps. On the other hand, fuzzy logic algorithms have also been employed to assess the spatial distribution of landslides. In addition, other algorithms such as Support Vector Machines (SVM) (Vapnik, 1995), Logistic Regression (LR) (Cabrera, 1994), Fisher's Linear Discriminant

* Corresponding author. Department of Civil Engineering, Gujarat Technological University, Nr.Visat Three Roads, Visat, Gandhinagar Highway, Chandkheda, Ahmedabad, 382424, Gujarat, India.
E-mail addresses: phambinhgtvt@gmail.com, binhpt@utt.edu.vn (B.T. Pham).

http://dx.doi.org/10.1016/j.envsoft.2016.07.005
1364-8152/© 2016 Elsevier Ltd. All rights reserved.

Analysis (FLDA) (Fisher, 1936), Bayesian Network (BN) (Friedman et al., 1997), and Naïve Bayes (NB) (Soria et al., 2011) are known to be applicable for binary classification problems that could be used for landslide susceptibility assessment.

Literature review shows that SVM has been applied efficiently and widely in landslide prediction (Kavzoglu et al., 2014; Peng et al., 2014; Pourghasemi et al., 2013). Marjanović et al. (2011) has carried out the comparison study of SVM with other methods, and concluded that SVM outperforms decision tree, logistic regression, and analytical hierarchy process in producing landslide susceptibility map. In another landslide study, Kavzoglu et al. (2014) also stated that the performance of SVM is better than the conventional logistic regression. In addition, Tien Bui et al. (2012a) reported that results derived from SVM can produce better landslide susceptibility map compared to decision tree, and naïve bayes methods.

LR has also been applied widely in landslide studies (Van Den Eeckhaut et al., 2006a; Yesilnacar and Topal, 2005). Das et al. (2010) proved that LR is a promising method for spatial prediction of landslides. Likewise, Akgun (2012) stated that LR has the best performance compared with other methods namely likelihood ratio, and multi-criteria decision analysis for landslide susceptibility map. In another comparison study, LR outperforms artificial neural network and likelihood ratio methods for landslide susceptibility analysis (Lee et al., 2007).

FLDA is one of the most used methods for complex data classification because it is simple enough to be tractable to itemized formal analysis in the projected area (Durrant and Kabán, 2010). FLDA has been applied as an efficient classifier in many fields such as pattern recognition (Cooke, 2002), medicine (Ambroise and McLachlan, 2002; Asamoah-Boaheng, 2014; Dudoit et al., 2002). For landslide problems, FLDA has been applied in very few studies (Murillo-García and Alcántara-Ayala, 2015; Rossi et al., 2010).

BN is considered as a promising method for hazard assessment (Liang et al., 2012). However, it is still rarely applied for the assessment of landslide hazard. Song et al. (2012a) used this method to assess susceptibility of earthquake-induced landslides, and stated that it shows high probability of landslide detection capability and it is a good alternative tool for landslide prediction. Liang et al. (2012) also stated that BN is useful for assessment of debris flow hazard.

NB method has also been applied successfully in some of the landslide assessment studies (Tien Bui et al., 2012a; Venkatesan et al., 2013). Pham et al. (2015b) applied this method for spatial prediction of landslides, and stated that it is an efficient machine learning method for landslide susceptibility assessment.

Overall, each above method has been applied successfully and efficiently for solving many real problems in many individual studies. Out of these methods, SVM and LR have been applied widely for landslide susceptibility assessment. However, FLDA, BN, and NB methods have been applied rarely for landslide prediction. Moreover, their performance has not also been compared in literature. Therefore, main objective of the present study is to evaluate and compare the performance of these machine learning methods for landslide susceptibility assessment. For this, part of Uttarakhand area of India which is prone to frequent landslides has been selected for landslide modeling. In this study, Receiver Operating Characteristic (ROC) curve and statistic index-based evaluations methods have been adopted to evaluate and compare landslide models.

2. Description of the study area

The study area is located between longitudes 78°37′40″E to 79°00′50″E and latitudes 30°23′15″N to 30°03′58″N in part of the Uttarakhand state of India, covering about 1325 km^2 (Fig 1). This area is situated in a subtropical monsoon region having annual average rainfall of 600 mm. Heavy rainfalls usually take place in monsoon season (June to September). Temperature during the year ranges from below 0 °C during winter to 45 °C during summer. Mean relative humidity varies from 25% to 85%.

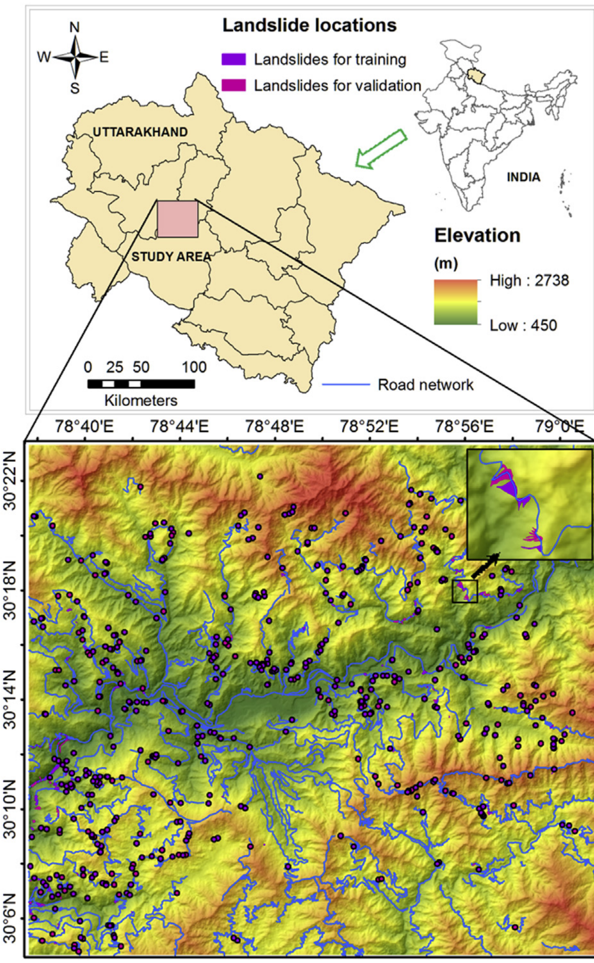

Fig. 1. Location of landslides in the study area.

Topographically, the study area belongs to mountainous region with intervening valleys. Elevation in the area varies from 450 m to 2738 m above sea level. Slope in this area is steep having slope angles up to 70°. Geologically, the area is structurally complex having folded igneous, sedimentary and metamorphic rocks (Agarwal and Kumar, 1973). Metamorphic rocks of Jaunsar group (phyllite and quartzite) are occupying major part of the area. Predominant soil types in the area are silty and loamy.

3. Methodology

Landslide susceptibility analysis in the present study has been carried out in six main steps (Fig 2): (1) preparing the geospatial database, (2) using feature selection method to select suitable landslide affecting factors for landslide analysis, (3) preparing training and testing data sets, (4) constructing landslide models, (5) validating and comparing landslide models, (6) developing landslide susceptibility map (LSM).

3.1. Preparation of geospatial database

Landslide inventory contains very important and indispensable

Download English Version:

https://daneshyari.com/en/article/6962342

Download Persian Version:

https://daneshyari.com/article/6962342

Daneshyari.com