Assessing the effectiveness of massive slope database in determining slope stability in Korea

Ivy F. Guevarra[1)]* · SangGi Hwang[1)] · TienChung Ho[1)] · ByungOk Yu[2)]

Problem:

The need to assure the safety of highway slopes rest in the hands of construction authorities and because of this, much had been spent to avert future slope failures. As more slopes are built, the cost of gathering relevant slope data becomes high and finding informative and significant pattern in massive amount of data becomes harder.

Solution and Objective of this study:

In response to this problem, this study will attempt to effectively mine the significant slope stability patterns from massive amount of data and determine the capacity of the method to correctly predict slope failure.

Method:

To extract patterns effectively, this study uses decision tree algorithm. Decision tree can extract patterns automatically from the database[1,2,3] (Witten and Frank, 2005). In context, the decision tree will classify the records whether it will fail or not. The classification is done using 6828 slope observation and 13 attributes. These attributes are lithology, tectonic domain, rainfall, seepage, slope dip, fracture orientation, notches, topography and weathering, height, length and natural slope.

IN order to find the important attributes in the database, individual attributes were ranked. Using the same set of data, attribute selectors were applied. Attribute selectors (Information Gain, ReliefF, Gain Ratio, Chi Square and Symmetrical Uncertainty)[4] automatically ranks the attribute. The result is presented in the Figure 1.

Result and discussion

The decision tree classified all slope instances in 180 rules. In summary, these rules suggest the that most influential slope attributes are seepage, slope dip, orientation of major fractures in slope and rainfall. The combination of these attributes will create rules that can effectively model the slope failure characteristics of the database. Other factors are significant however their effect is restricted in lower branches only. These attributes are notches, domain, lithology, topography and weathering. The rest have little effect in predicting slope failure.

Keywords: decision tree, database, slope stability

1) PaiChai University, Department of Civil, Environmental and Railroad Engineering (sghmap@pcu.ac.kr)

2) Geotechnical Research Division, Highway Research Center, Korea Highway Corporation

In comparison, important attributes from the decision tree both found in higher and lower branches of the model were predicted by the individual ranking from attribute selection. This ranking considered not only fracture orientation, seepage rainfall and slopedip but also lithology.

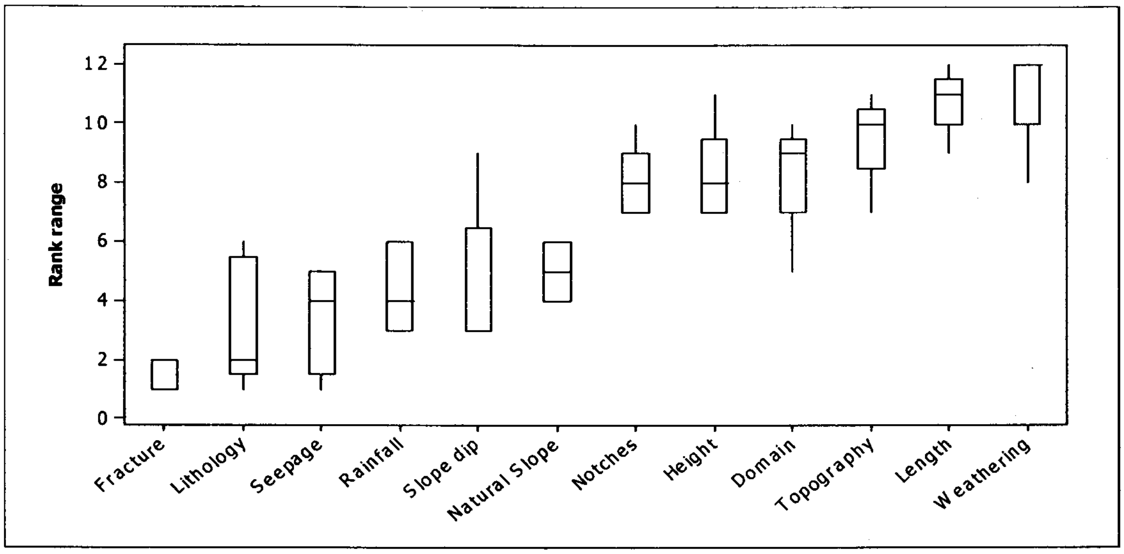

Figure 1: Ranking of the importance of individual attributes in terms of predicting future failure.

By assessing the predictive accuracy of the output model, applying decision tree turns out to be successful in mining data from a massive database. The rules correctly predicted 71.98% of the original data. Further, if we restrict our assessment in failed slopes, the decision tree can predict up to 80.3 % of all the slopes that failed in the past. At this prediction rate, the cost of building massive databases are justified.

Table 2: Evaluation result on the database.

Number of instances	100%	6828
Correctly predicted instances	71.98%	4915
Incorrectly predicted instances	28.02%	1913

Table 3: Detailed accuracy by class of the database.

Class	TP rate (%)	FP rate (%)	Precisi on (%)	Recall (%)
Safe	60.6	19.7	69.4	64.7
Fail	80.3	39.4	73.5	76.8

6. References

[1] Korean Weather Bureau (www.kma.go.kr)

[2] Ministry of Construction and Transportation, 2004, Route Designation Status of National Highway (www.moct.go.kr)

[3] Korean Highway corporation

Witten, I. and E. Frank, 2005, Data Mining Practical Machine Tools and Techniques. Elsevier, San Francisco.

[4]WEKA software version 3.4. University of Waikato (www.cs.waikato.ac.nz/ml/weka)