Research on and application of tunnel structure defects prediction using machine learning methods

Bo Shi[1,2], Hui Su[1], Xu Du[1], Bao Jiao[1] and Lin Wang[*1,3]

[1] SGIDI Engineering Consulting (Group) Co., Ltd, Shanghai 200093, P.R. China

[2] School of Computer Science and Technology, Fudan University, Shanghai 2000438, P.R. China

[3] Department of Geotechnical Engineering, College of Civil Engineering, Tongji University, Shanghai 200092, P.R.China

[*] Email: 1410024@tongji.edu.cn

ID: CEE 1467

INTRODUCTION

Previous studies on tunnel assessments mostly used data mining technology to evaluate the state of tunnel structure and put forward corresponding measures (Adoko A. et al.,2013;Mahdevari S. et al.,2013) and they concerned with the deformation prediction of the ring segments such as convergence. Unlike predicting the deformation in tunnels in previous studies, the paper first provided an idea for predicting defects and defect types using machine learning methods. Accurately and efficiently predicting possible tunnel diseases has important engineering guiding significance to help tunnel maintenance departments determine the inspection scope and avoid major disasters. In this study, three machine learning methods, the decision tree, random forest, and XGBoost, were applied to the prediction of tunnel structure diseases to help the tunnel maintenance department understand the development trend of the disease and make timely decisions on prevention and control measures.

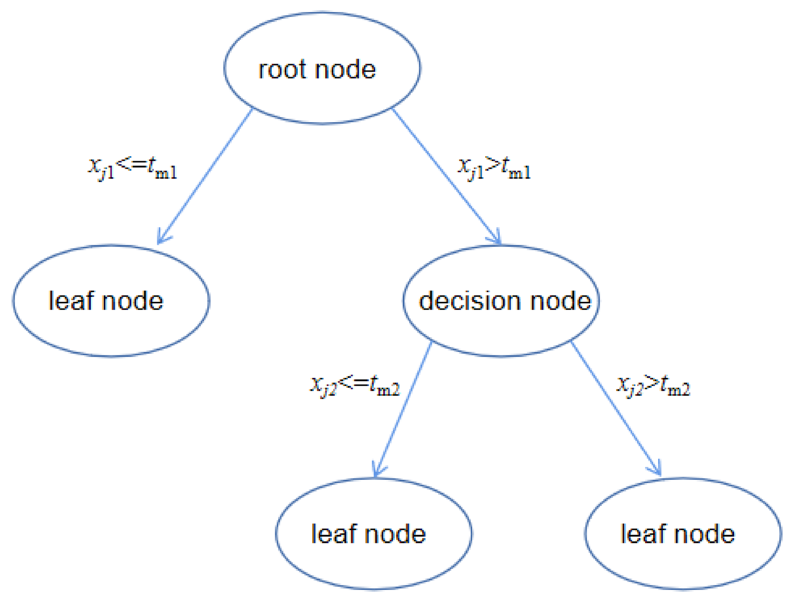

Figure 1. Sketch map of a decision tree

METHODOLOGY

Decision tree (Quinlan, 1986) is a machine learning prediction model that represents a mapping relationship between features and target values . The sketch map of a decision tree is shown in Figure 1. For a given dataset, it can be recursively split into 'tree' as the sketch map.Random forest (Breiman, 2001) is an ensemble method that combines several base estimators to improve accuracy and stability over a single estimator. XGBoost (Chen, 2016) is also an ensemble method that combines several base estimators. In this study, CART(Breiman, 1984) was used as the base estimator, and "Gini" was used as the loss function.

RESULTS

In this study, data of 68,055 segment lining rings of six subway lines in a city were fed into the machine learning models mentioned above. According to defects records from 2014 to 2016 and corresponding convergence and characteristic data, defects conditions in 2017 were predicted and compared with real defects conditions in 2017 to analyze the prediction effect of the models. We used four indexes, accuracy, precision, recall, and F1 values, to comprehensively evaluate the effect of the model calculated in Table 1. We compared the real condition of tunnel segment defects with the predicted condition to generate the confusion matrix (Figures 2–4), which represents the effect of the models. We also plotted the ROC curves of the three models and the corresponding AUC values were calculated in the Figure 5.

Table 1. Evaluation indexes of the machine learning models

Algorithms	Accuracy	Precision	Recall	F1 values
Decision tree	93%	98.4%	74%	83%
Random forest	93%	98.4%	69.5%	81%
XGBoost	93%	99.8%	69%	81.4%

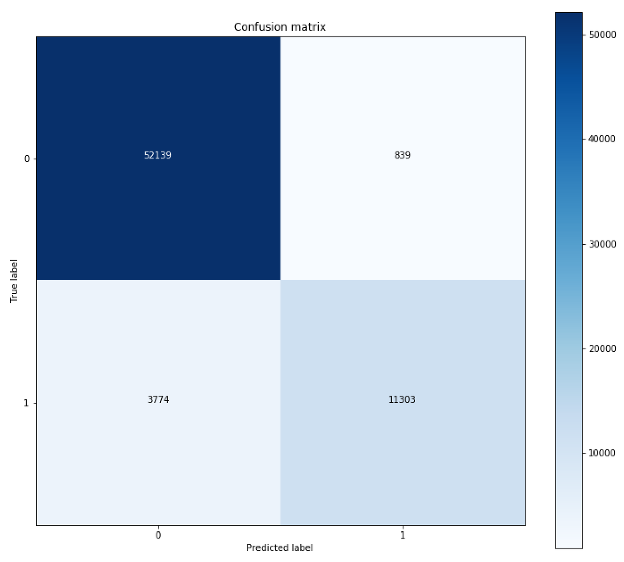

Figure 2. Prediction results of Decision Tree model of binary classification

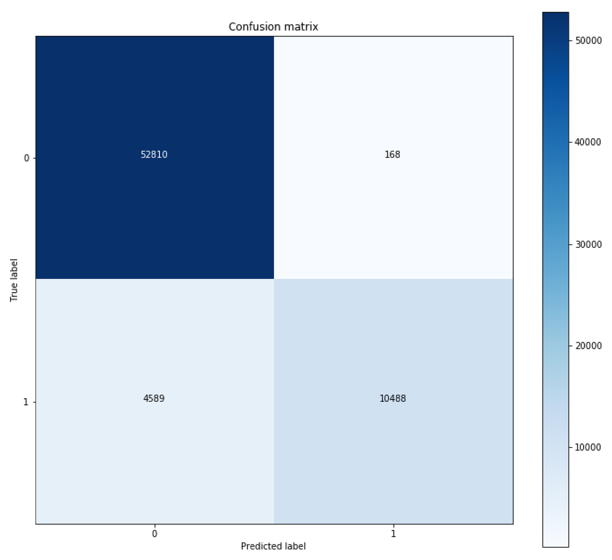

Figure 3. Prediction results of Random Forest model of binary classification

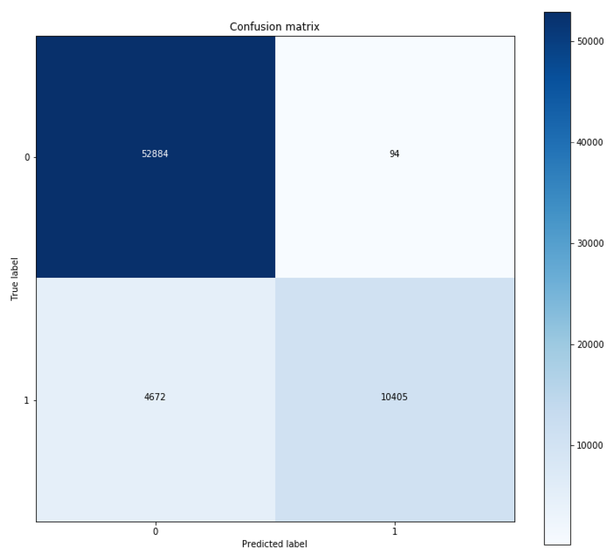

Figure 4. Prediction results of the XGBoost model of binary classification

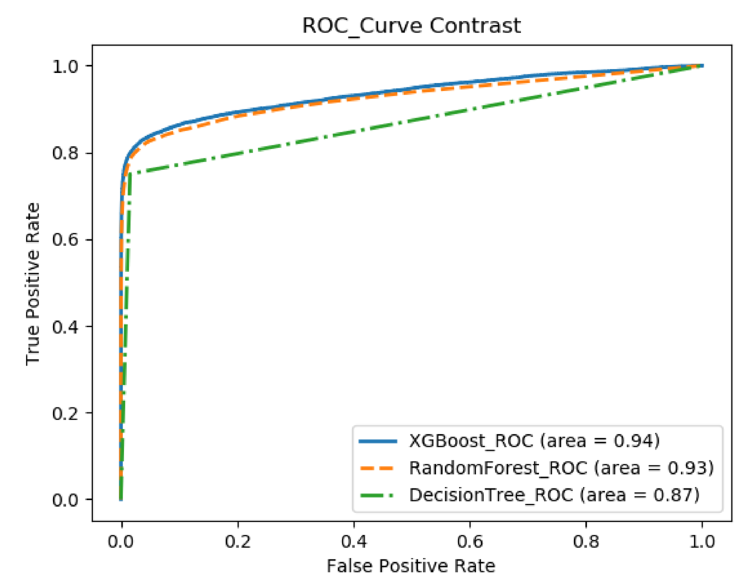

Figure 5. ROC curve and AUC values of the three models

CONCLUSION

Through the three models, we predicted tunnel structural defects and obtained satisfactory results. The accuracy rates were all above 90%, and the F1 values were all above 80%.The AUC values of XGBoost model, Random Forest model, Decision Tree model are 0.94,0.93 and 0.87 respectively which means that the algorithms we chose are robust for imbalanced engineering data. Overall, our proposed method can be useful to aid tunnel maintenance departments owing to the high prediction accuracy rate.The prediction results can provide auxiliary decision-making assistance for tunnel maintenance departments and relevant government regulatory departments to prevent and control tunnel structural diseases and focus on tunnel sections where serious diseases may occur so as to further clarify the development trend of tunnel diseases.

REF

Adoko A.,Jiao Y.,Wu L., et al..2013.Predicting tunnel convergence using Multivariate Adaptive Regression Spline and Artificial Neural Network.Tunnelling and Underground Space Technology, 38:368-376.

Breiman.2001.Random forests.MachineLearning,45(1):5-32.

Breiman,J. Friedman, R. Olshen, and C. Stone.1984. Classification and Regression Trees. Wadsworth, Belmont, CA.

Chen T.,Carlos Guestrin.2016. XGBoost:A Scalable Tree Boosting System. SIGKDD, doi:10.1145/2939672.2939785

Mahdevari S.,Hamid Shirzad Haghighat, Seyed Rahman Torabi.2013. A dynamically approach based on SVM algorithm for prediction of tunnel convergence during excavation. Tunnelling and Underground Space Technology,38:59-68.

Quinlan J.R.1986. Introduction of decision trees. Machine Learning,1(1):81-106.

ACKNOWLEDGEMENTS

The authors would thank Social Development Projects of Shanghai 'Science and Technology Innovation Action Plan'(18DZ1205900,19DZ1200801) and National Natural Science Foundation of China (42002272) for support.