Prediction of Accident Types and Accident-Causing Objects Using

3.2 Model Optimization

In this study, two separate models were developed to predict dangerous objects and accident types. The optimal algorithm for each model was determined by comparing the performances of the four algorithms. Those algorithms are the most commonly accepted in the latest research on the subject. The four algorithms were tuned using training data sampled via a grid search. Five-fold cross-validation was performed to determine the optimal hyperparameter combination using the grid-search best-param libraries. The hyperparameter combinations for each algorithm are explained in Table 3.

Table 3: Grid search value for algorithms

Algorithm	**Grid search value (hyperparameter)**
XGBoost	Max_depth: 7-11 (interval: 2), Min_child_weight: 1-7 (interval: 2), colsample_bytree: 0.25-0.75 (interval: 0.25), n_estimators: 10-300 (interval: 10)
CatBoost	Max_depth: 7-11 (interval: 2), learning_rate: 0.01-0.5 (interval: 0.1), n_estimators: 10-300 (interval: 10)
LGBM	Max_depth: 7-11 (interval: 2), Num_leaves: 7-13 (interval: 2), learning_rate: 0.01-0.5 (interval: 0.1), n_estimators: 10-300 (interval: 10)
RF	Min_samples_split: 3-7 (interval: 2), Min_samples_leaf: 3-7 (interval: 2), n_estimators: 10-300 (interval: 10)

After training the model with the optimal hyperparameters, an optimal algorithm was selected by

c

Figure 2: Results of the model performance comparison

4.2 Model Evaluation Results and Discussion

Table 5 summarizes the accuracy, precision, recall, and F1-score of each model. Both models performed acceptable, with the accident-object prediction model performing better with an accuracy of 0.879, precision of 0.873, recall of 0.878, and F1-score of 0.874. The4Qq0.50.873

development of a model to evaluate the safety level of construction sites and the impact analysis of accident-causing factors on site safety.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1A2C2003696) and also supported by the National R&D Project for Smart Construction Technology funded by the Korea Agency for Infrastructure Technology Advancement under the Ministry of Land, Infrastructure and Transport (No. 22SMIP-A158708-03). This work was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2023-20241758).

References

Chang, T., Lee, G., & Chi, S. (2023). Development of an Optimized Condition Estimation Model for Bridge Components Using Data-Driven Approaches. *Journal of Performance of Constructed Facilities*, *37*(3). https://doi.org/10.1061/JPCFEV.CFENG-4359

Cheng, C. W., Leu, S. Sen, Cheng, Y. M., Wu, T. C., & Lin, C. C. (2012). Applying data mining techniques to explore factors contributing to occupational injuries in Taiwan's construction industry. *Accident Analysis and Prevention*, *48*