일반연구보고서 2022-5

빅데이터 기반 건축물 화재 예측 모델 개발 연구

A Study on the Development of Building Fire Prediction Model Based on Big Data Analysis

조영진 Cho, Youngjin
허한결 Heo, Hankyul
안의순 Ahn, Euisoon
류수연 Ryu, Suyoon
송유미 Song, Yumi
현태환 Hyeon, Taehwan

A Study on the Development of Building Fire Prediction Model Based on Big-Data Analysis

Summary

Cho, Young Jin
Heo, Han Kyul
Ahn, Eui Soon
Ryu, Su Yeon
Song, Yu Mi
Hyeon, Tae Hwan

Safety accidents in buildings due to natural disasters such as earthquakes, floods, typhoons, and social disasters such as fires and crimes are increasing, and concerns about the large–scale human and property damage are also increasing. According to the "2020 Disaster Yearbook" published by the Ministry of Public Administration and Security, the amount of damage to buildings caused by typhoons, heavy rains and earthquakes between 2011 and 2020 was very high. In addition, there is a concern about large–scale human and property damage in the event of a safety accident due to an increase in medium and large–sized buildings. According to the Construction Safety Management Comprehensive Information Network, a total of 34,385 safety accidents occurred at construction sites in 2021, and a number of casualties from fires and collapses are continuously reported. The increase in aging buildings is accompanied by decrepit mechanical and electrical facilities of buildings, which is a factor that increases the risk of fire, so countermeasures are needed.

Accordingly, there is an increasing need for research on the use of data in architectural and urban spaces to solve disasters and disaster problems in buildings. Fire countermeasures for existing buildings are mainly to check fire prevention facilities to minimize damage in the event of a fire, and firefighters request the use of building drawing information in the electronic architectural administration information system to respond to fires. In light of the increasing amount of information on various and vast architectural and urban spaces such as the electronic architectural administration information system and the national fire information system, research is needed to establish a disaster and disaster prediction system based on big data analysis. With the development of fourth industrial revolution technologies such as AI, machine learning, and deep learning, the accuracy of predicting disasters and disasters is increasing significantly, and problem–solving skills using architectural and urban space data are secured, so research is needed to apply these skills.

Therefore, this study aims to establish a methodology for predicting dangerous areas to prevent safety accidents in buildings using attribute information of buildings and fire occurrence data in buildings, and to propose a model for predicting disaster and disaster risk areas. First of all, we intend to derive a linkage plan between fire occurrence data and electronic architectural administration information system data, and analyze it to establish a methodology for predicting fire occurrence risk areas. In particular, it is intended to derive a plan to link building fire occurrence data with electronic architectural administration information system data. In addition, we would like to propose a disaster and disaster risk area prediction model. Based on the fire data risk area prediction methodology, we would like to propose a model that can predict disaster and disaster risk areas occurring in buildings and spaces such as collapse, heat wave, and flooding, and suggest a direction that can be used in policy.

Chapter 2 examines domestic and foreign disaster big data systems, big data analysis methodologies, and related research trends.

The state has established and operated various crisis management systems for disaster prevention, management, and recovery, and related systems are operated based on various disaster and disaster data. The disaster prevention and response ministries in Korea are overseen by the Disaster Safety Management Headquarters of the Ministry of Public Administration and Security, and operate the National Disaster Management Information System and the Integrated Safety Information Disclosure System to prevent and respond to disasters.

Fire and crime, which are social disasters, have established and operated individual systems. Fire is used as data for establishing and responding to fire prevention policies by providing information on the risk of fire occurrence by region through the National Fire Information System. The geographical profiling system operated by the National Police Agency mainly performs crime risk measurement for crime prevention. Overseas fire occurrence prediction systems primarily focus on fire occurrence and spread prediction systems through large-scale fire occurrence monitoring. Examples include FIRECAST and Fire Danger Forecast, which is operated by the U.S. Department of Interior.

The implications derived through the review of big data prediction systems related to disasters and disasters at home and abroad are that it is necessary to establish a system for collecting, processing, and analyzing related big data to prevent disasters and disasters. In addition, in order to respond quickly to disasters, take measures, and minimize damage, related data should be provided easily to policymakers and the general public, and building attribute information is important to establish a disaster and disaster prediction model for buildings.

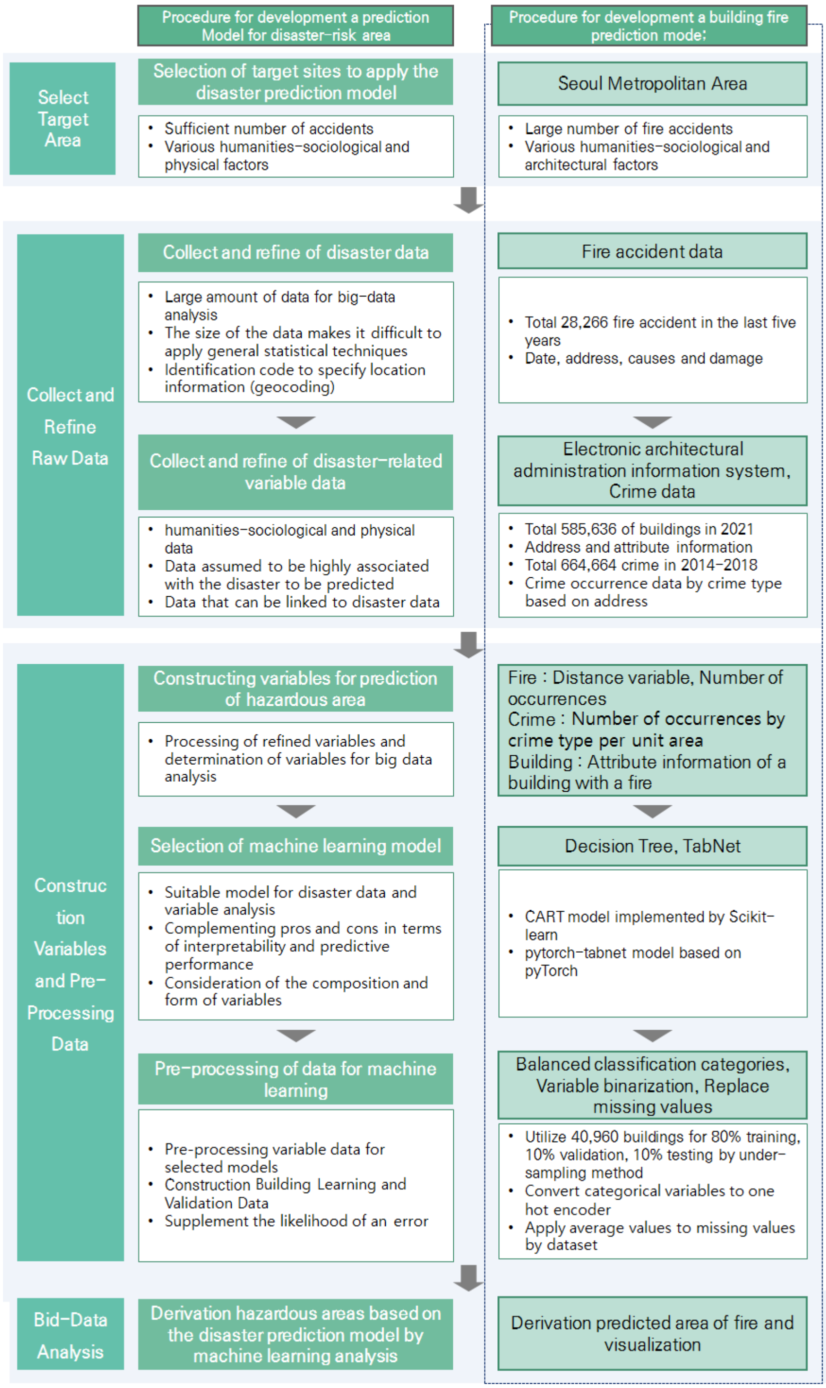
Looking at big data-based disaster and fire prediction research trends, the focus is on the types and lists of data that should be used for prediction and the big data research methodology that utilizes existing data. The type of data, the degree of construction, what data can be used other than disaster and disaster data, and the focus is on research on scenario construction for prediction.

In Chapter 3, related data selection and linkage measures were prepared to develop a building fire prediction methodology, and a building fire risk area prediction methodology was established.

A place with general characteristics of the city was selected so that it could be applied to various spatial units as a matter to consider when selecting a target site for the establishment of a prediction model for a fire hazard area. Therefore, in this study, Seoul was the target. Next, in order to apply to the big data analysis methodology, the fire data to be analyzed in this study was linked to the electronic architectural administration information system data. First of all, the structure of the fire occurrence data and the construction administration data was checked, and the PNU code containing address information was generated with the linkage KEY of the two data to perform the linkage.

After linking fire data and electronic architectural administration information system data, spatial informatization, or geocoding, was performed because distance variables will be introduced when analyzing big data. Building integration information was used as Basemap, and in the case of crime data, the data were enormous, so instead of extracting crime-prone buildings, it was derived in units of aggregated districts. As the next step, the ArcGis program was used to build a dataset for big data analysis, and three types of fire-related variables, six types of crime-related variables, and six types of building-related variables were built as independent variables, and whether a building fire occurred was built as dependent variables.

After setting variables for big data analysis, a machine learning model was selected. In this study, the CART model implemented by scikit-learn was used in consideration of the fact that fire was a binary variable. In addition, the TabNet model used in this study is a pyTorch-tabnet model based on pyTorch, and a classification model was used like the decision tree. After that, a total of 40,960 data were extracted as final analysis data by pre-processing and under-sampling according to each model, and 80% of the data to be analyzed was used as training data, 10% as verification data, and 10% as test data. In this way, the development of a prediction model for disaster and disaster risk areas was carried out in the order of data processing for analysis, variable construction, and big data analysis model selection to suit the model.



[Figure] Procedure for development prediction model for disaster-risk area

Source : Inhouse

In Chapter 4, the prediction model of buildings in dangerous areas was applied based on the construction of the final dataset and the selection of machine learning models performed in Chapter 3.

First of all, the current status of fire in buildings in Seoul was identified through technical statistics on building fire data. In the case of Seoul, the number of fires in apartments and detached houses was high, and in the case of total area, there were many fires in buildings exceeding 200㎡ and 500㎡ or less, and in the fire statistics by floor, there were many fires in low floors below five stories. In addition, as the aging progressed, the number of fires increased.

In big data analysis, the TabNet model was used to predict fire-prone buildings among decision tree models and deep learning models, and the types of fire-prone buildings and fire risk influencing factors were extracted through analysis of machine learning models. Through this process, the prediction model was verified and the characteristics of buildings vulnerable to fire were interpreted. As a result of the analysis through the decision tree model, among buildings with a total floor area of 5,160㎡ or more, the fire incidence rate in types excluding educational research facilities was the highest at 12.0%.

Through the extraction of fire risk influencing factors derived from the TabNet model, the possibility of using the predictive model through analysis of the deep learning model was verified. As a result of the analysis, the accuracy for the verification dataset was 67.5%, and the accuracy for the test dataset that was not used for learning and early interruption determination was 65.8%. As a result of extracting the factors influencing fire risk, the total area was the most relevant among the building properties, and the use was the next highest. It was found that the relationship between the violation building and the approval for use (oldness) was relatively low. In the case of crimes, the number of five major crimes was highly related, while the number of individual crimes was low. In the case of the distance variable, the radius of 1km of the fire area was higher than that of the surrounding 1㎢ and 500m.

The machine learning-based building disaster and disaster prediction model proposed in this study is a methodology that can identify fire vulnerability prediction, the type of vulnerability, and the cause of vulnerability judgment for each building unit. As the last step of the analysis, the fire vulnerability prediction result was visualized based on spatial information, and the applicability of GIS visualization technology was reviewed.

Finally, the limitations of this study are that the analysis of all data was not performed due to some errors in fire occurrence data and electronic architectural administration information system data, and the prediction accuracy is relatively low because the types of variables applied in big data analysis are not diverse.

Therefore, in order to improve and upgrade the accuracy of the big data-based safety accident prediction model developed in this study, it is necessary to perform tasks to increase the integrity of related data and apply various humanities and sociological variables to increase the prediction accuracy.

Keywords :

Big-Data, Building Fire, Decision Tree, TabNet, Machine Learning