Research Article

pISSN : 2287-8920
https://doi.org/10.22716/sckt.2021.9.3.031

화재 조사 데이터 분류별 기상정보 기반 화재 위험도 예측

송동우[1]
[1]한라대학교 소방방재학과 조교수

Fire-Risk Prediction Based on Weather Information by Classification of Fire Investigation Information

Dongwoo Song[1]
[1]Assistant Professor, Department of Fire and Disaster Prevention Engineering, Halla University

[1]**Corresponding author:** dw.song@halla.ac.kr

Received August 10, 2021; **Revised** September 12, 2021; **Accepted** September 22, 2021

ABSTRACT

화재 안전을 위해서, 화재 발생을 방지하는 것이 가장 근본적인 방법이다. 기상의 변화는 인간의 생활의 변화와 밀접한 관계가 있고, 기상 변화에 따른 화재 발생 예측은 화재 예방을 위해 많은 기여를 할 수 있을 것이다 . 본 연구에서는 예측 기술을 활용하여 기상의 변화에 따른 화재 발생의 변화를 분석하고자 한다. 예측 모델은 온도, 습도, 풍속, 지역의 4가지 입력 변수를 기반으로 하였다. 이러한 기상 조건을 기반으로 화재 유형, 발화원, 발화 원인 및 발화 물질과 같은 여러 화재 특성에 대한 화재 위험 예측 정확도를 확인하였다. 분석을 위해 '인구 시간당 화재 발생 빈도'를 산출하고 이를 위험등급으로 구분하는 데이터 전처리를 하였고, 인공신경망과 의사결정트리를 활용한 화재 예측 모형을 만들었다. 분석 결과 세부 변수별 화재 위험도의 예측 정확도를 비교하였고 기상과 연관성이 높은 특정 유형의 화재를 확인하였다. 본 연구는 기상 조건에 따라 화재 발생 위험이 높은 세부 특성을 파악하고 추후 화재 특성별 위험도를 도출하기 위한 기초자료로 활용 가능할 것이다.

For fire safety, preventing the occurrence of a fire is the most fundamental method. Changes in weather are closely related to changes in human life, and prediction of fire occurrence according to changes in weather can make a great contribution to fire prevention. This study intends to analyze the change in fire occurrence according to the change of weather by using the prediction technology. The prediction model's results were based on 4 input variables (temperature, humidity, wind velocity, and region); it measured the fire-risk prediction accuracy for several fire characteristics such as fire type, ignition source, ignition cause, and ignition material based on those weather conditions. For the analysis, data pre-processing was performed to calculate the 'population frequency of fires per hour' and classify them into risk grades, and the fire prediction model using artificial neural networks and decision trees was created. As a result of the analysis, the prediction accuracy of the fire risk for each detailed variable was compared, and a specific type of fire highly related to the weather was identified. This study can be used as basic data to identify detailed fire characteristics with a high risk of fire depending on weather conditions and to derive the risk level for each fire characteristic in the future.

Keywords: Fire safety, Weather conditions, Risk prediction, Fire type, Source of ignition, Cause of ignition, Ignition material

1. Introduction

In studies of risk, prevention is the most important stage. For that reason, significant efforts have been made in many fields to predict risks and thus prevent various negative outcomes. Much prediction research has been conducted, including predictions of disease in the medical field; predictions of mechanical and facility accidents in the industrial sector; and predictions of earthquakes, landslides, and fires in disaster studies.

Although researchers have conducted many studies on data-based fire prediction, relatively few of those were on the relationship between weather and fire risks. Among the prior work on the relationship between weather conditions and fires other than woodland fires, Yang et al. concluded that air temperature has a low correlation with dwelling fires[1]. However, some scholars have shown a possible correlation between fire-related factors and weather factors. For instance, weather conditions such as extreme cold or heat can create conditions that carry greater fire risk[2]. Holmes discussed the change of the occurrence rate for chimney fire according to temperature and other weather changes using statistics and data-mining techniques, revealing a trend that chimney fires generally increased as the temperature dropped (indicating a seasonal influence); the temperature showed a somewhat nonlinear relationship[3]. Chandler observed a significant relationship between the number of fires and the daily low temperature, reporting that bad weather conditions increased specific fire causes related to space heating, wires and cables, and smokers' materials[4]. Wang et al. found that Gross Domestic Product and humidity affected the fire situation in western China[4]. Corcoran et al. suggested that temperature extremes increase fire risk for all the investigated fire types[2]. Song et al. conducted a relational analysis of weather factors and fire-occurrence factors. Using the frequency of fire occurrence and various meteorological factors, the resulting relational graphs consider each fire characteristic, including source of ignition, cause of ignition, and ignition material. For humidity, the trend was that fires occurred more frequent as the humidity decreased. For temperature, various trends occurred according to the types of independent and dependent variables. One trend showed that fires occurred more often as the wind speed increased[5]. The present study was developed based on the research performed by Song et al., and it considers the accuracy of fire-risk predictions for factors such as the fire's type, cause, and materials, as well as the weather conditions.

Models were developed to predict fire risk using fire characteristics such as fire type, source of ignition, cause of ignition, and ignition material. This research study had three primary objectives: (1) to preprocess the available fire data and meteorological data in order to predict fire risk; (2) to develop fire-risk prediction models using two data-mining techniques: artificial neural networks (ANNs) and decision trees (DTs); and (3) to examine prediction accuracy for characteristics related to the fires and to thus evaluate the models.

2. Materials and Methods

The data used in this study are fire investigation data and meteorological data for 6 years from 2007. Temperature, humidity, and wind speed data for each region were used as meteorological data. For the fire incidents, data from the Korean National Emergency Management Agency's national fire-information system were used. This system collects all of the fire-investigation information from fire departments nationwide. This study analyzed the following fire characteristics: fire type, source of ignition, ignition material, and cause of ignition. Each of these fire characteristics (classified in Table 1) acted

as a dependent variable. Here, “fire type” refers to the property type. “Source of ignition” refers to the source of the flame or heat that caused the fire. “Cause of ignition” refers to the factor that influenced the fire’s ignition. Finally, “ignition material” is the material that first ignited and that burned until the fire eventually became hard to control.

Table 1. Classification of the fire characteristics in the fire-investigation data

	Categories
Fire type	dwelling, non-dwelling, forest or field, automobile or train, miscellaneous (waste material, etc.)
Source of ignition	cigarette or lighter; friction, conduction, or radiation; flame or spark; spontaneous combustion; actuated equipment; explosive or firework; chemical combustion; miscellaneous; unknown
Cause of ignition	gas leak or explosion, traffic accident, mechanical factors, carelessness, natural factors, electrical factors, chemical factors, miscellaneous, unknown
Ignition material	furniture; flammable gas; signboard or awning; food; waste material; hazardous material; automobile, train, ship, or aircraft; electrical or electronic equipment; paper, wood, or hay; bedding or textiles; synthetic resins; miscellaneous; unknown

Data-mining techniques were applied to extract the desired information from saved data or to find patterns and/or relationships included in the data. Several algorithms were proposed, depending on the purpose and function of the prediction. Multiple-regression models have been widely used because they are easy to interpret. However, it can only be used for linear mapping. Alternatives such as DTs, ANNs, and random forests have emerged to address these drawbacks[8]. DTs and ANNs have been used in various fields. DTs have been successfully applied as high-dimensional classifiers in studies in a wide variety of fields[9-14]. ANNs are widely used data-mining tools that are based on nonlinear functions[10]. ANNs can detect patterns that are too complex for other techniques to identify[13]. ANNs have exhibited proven performance in many fields[9-13,15,16]. For this study, ANN and DT models were selected as the data-mining techniques and a feed-forward neural network with one hidden layer was used. The software used for analysis is SPSS Modeler. The schematic diagram for data analysis is shown in Fig. 1.

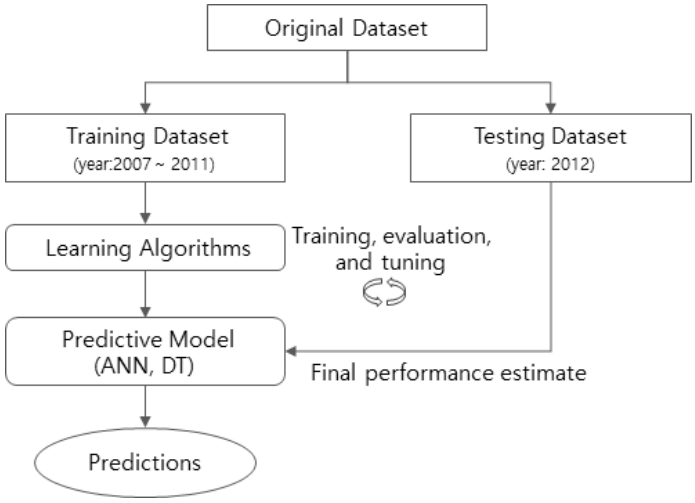

Fig. 1. Schematic diagram for data analysis

Raw data were processed for analysis. For raw data, scenarios were created using region, time slot, temperature, humidity, and wind speed. A scenario is a combination of data values and input variables. Each scenario is presented as a frequency of fires per 100,000 people. All dependent variables were divided into 5 levels based on fire risk. Classification methods are classified into 5 levels according to the quantile classification method[6,7]. The following data process is performed for all 38 dependent variables for all fires, including fire type, source of ignition, cause of ignition, and ignition material. Fig. 2 is an example of creating scenarios.

Scenario	Region	Temp.	Humidity	Wind velocity	fire frequency	Number of scenario case	population	Hourly fire frequency per 100,000 persons	Class
S1	Gangwon	t1	h1	wv1	0	1	1503806	0.0000	CLASS 1
S2	Gangwon	t1	h1	wv3	0	1	1503806	0.0000	CLASS 1
S3	Gangwon	t1	h2	wv2	2	9	1503806	0.0148	CLASS 4
S4	Gangwon	t1	h2	wv3	6	13	1503806	0.0307	CLASS 5
S5	Gangwon	t1	h2	wv4	1	2	1503806	0.0333	CLASS 5
S6	Gangwon	t1	h3	wv1	2	10	1503806	0.0133	CLASS 4
S7	Gangwon	t1	h3	wv3	0	4	1503806	0.0000	CLASS 1
S8	Gangwon	t1	h3	wv4	1	2	1503806	0.0333	CLASS 5
S9	Gangwon	t1	h4	wv1	14	49	1503806	0.0190	CLASS 4
S10	Gangwon	t1	h4	wv2	3	3	1503806	0.0665	CLASS 5
S11	Gangwon	t1	h5	wv1	31	168	1503806	0.0123	CLASS 4
S12	Gangwon	t1	h5	wv2	5	6	1503806	0.0554	CLASS 5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Fig. 2. Examples of data preprocessing

3. Results

3.1 Results for the fire types

For both prediction models, the results for the 4 input variables were evaluated for each fire type. The fire type is analyzed, with building-structure fires separated into dwelling and non-dwelling fires, using the prediction results from the test set.

Fig. 3 shows the total prediction accuracy for the test set analyzed. The comparison of the fire types shows that, for the ANN model, the total prediction accuracy was 46.33% for dwelling fires and 42.48% for non-dwelling fires. Regarding the prediction accuracy for dwelling fires at individual levels, level 3 had the lowest accuracy. The prediction accuracy for individual levels was similar for non-dwelling fires. The forest or field fires had high values for total prediction accuracy (69.60%). For miscellaneous (wastes and etc.) fires, the total prediction accuracy shows to be 51.61%. In terms of prediction accuracy for each level, in the miscellaneous fires and the forest or field fires, the prediction accuracy for level 3 was lower than those of the other levels. For the automobile or train fire type, the total prediction accuracy was 50.92% and the prediction accuracies for level 3 and level 4 were low. The DT model had similar results. The prediction accuracy for each level was also relatively low at level 3 and level 4. The DT model had similar results.

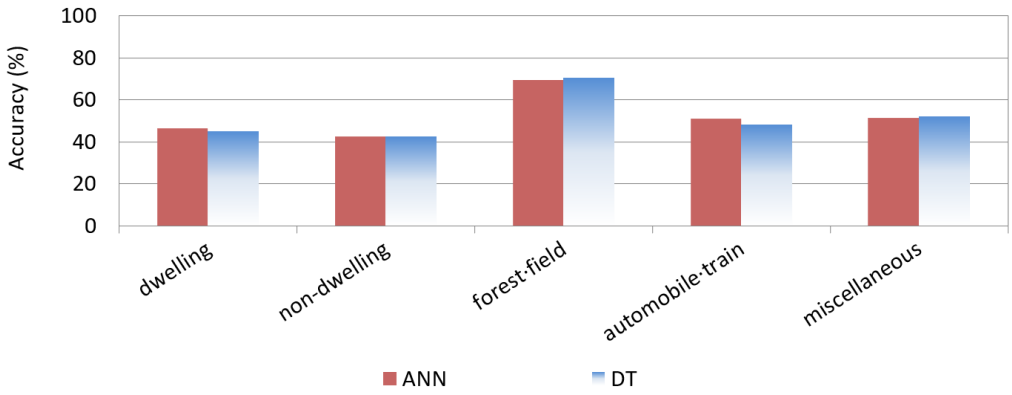

Fig. 3. Accuracy of the test set according to the ANN and DT prediction models for each fire type

3.2 Results for the sources of ignition

Comparing the ignition sources, the spontaneous combustion, explosive or firework, and chemical combustion sources had high total prediction accuracies but were hard to predict for individual levels. The cigarette or lighter; friction, conduction, or radiation; flame or spark, and unknown sources all had total prediction accuracies of approximately 50%. The prediction accuracies for each level were relatively low in the middle levels for these categories. For the miscellaneous category, the prediction accuracy for level 5 was very low. The actuated equipment category had total prediction accuracy of 40% or lower (which was the lowest of all the categories), but it also had the most uniform prediction accuracy across levels. Comparing the models' analysis results, as shown in Fig. 4, they had similar results for total accuracy.

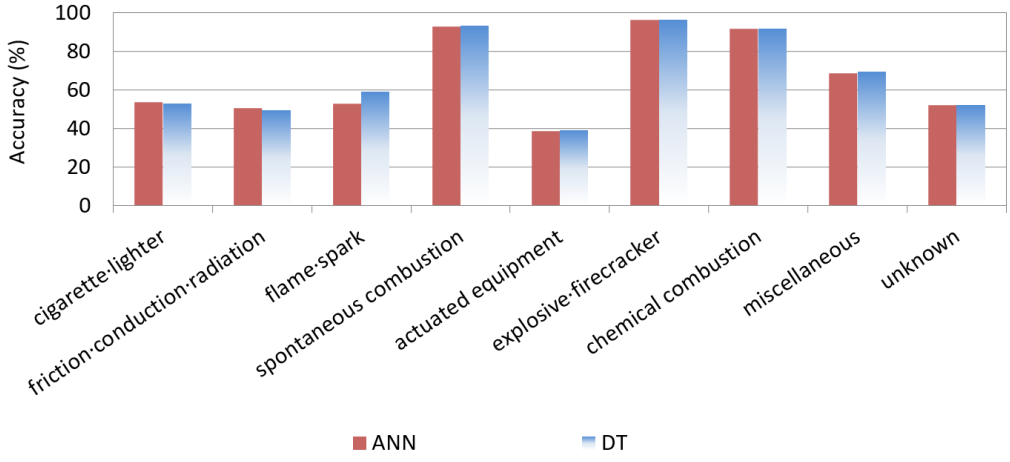

Fig. 4. Accuracy of the test set according to the ANN and DT prediction models for each source of ignition

3.3 Results for the causes of ignition

Comparing the results for the fire causes, for each fire cause except mechanical factor, negligence, and unknown, the fire characteristics were difficult to predict at each individual level. The mechanical factor and unknown categories had total prediction accuracies of approximately 50%, but they had different trends in the ANN and DT models. In the ANN model, the

prediction accuracies for the middle levels were relatively low, and in the DT model, levels 3 through 5 showed similar prediction accuracies. For the negligence cause, a high total prediction accuracy of greater than 70% was shown, and the prediction accuracy decreased as the level increased. Comparing each model's analysis results, as shown in Fig. 5, the models had slight differences in terms of total accuracy for each of the fire characteristics, but they had similar trends in most cases.

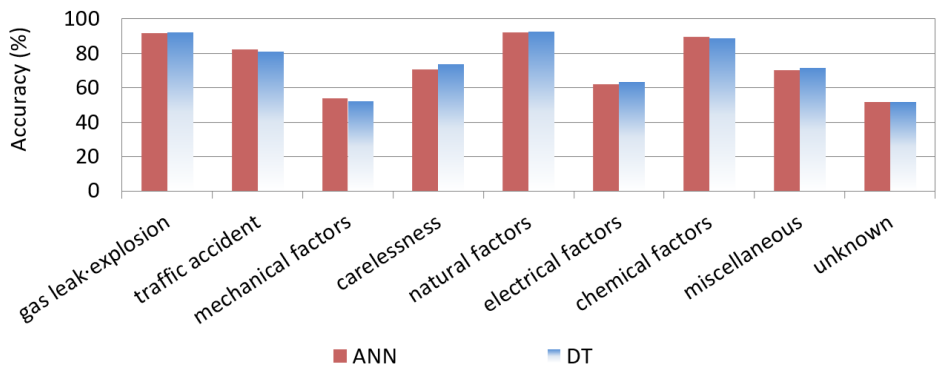

Fig. 5. Accuracy of the test set according to the ANN and DT prediction models for each cause of ignition

3.4 Results for the Ignition Materials

According to the comparison results for the ignition materials, the prediction at the individual levels was complicated for these categories: furniture; flammable gas; signboard or awning; unknown; hazardous material; automobile, train, ship, or aircraft; electric or electronic equipment; bedding or textiles; and miscellaneous. Total prediction accuracies were approximately 60% for the food category and approximately 50% for the categories of waste; paper, wood, or hay; and synthetic resin. In addition, the prediction accuracies in the middle levels were relatively low for these materials. For paper, wood, or hay, the prediction accuracy was higher than for other variables at each level. Comparing the models' analysis results, as shown in Fig. 6, the models had slight differences in terms of total accuracy for each of the fire characteristics, but there were similar trends in most cases.

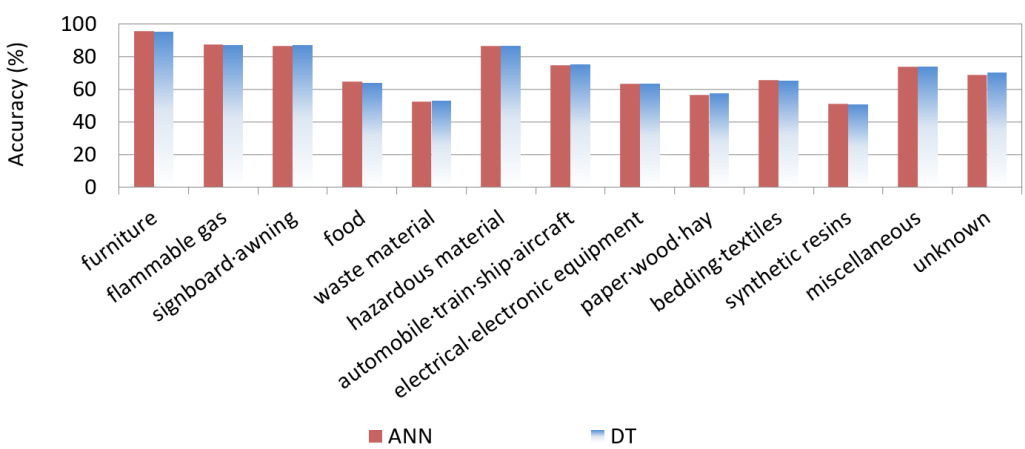

Fig. 6. Accuracy of the test set according to the ANN and DT prediction models for each ignition material

4. Discussion and Conclusions

The prediction model for each fire characteristic presented the results for 4 input variables. For the result of fire type, the dwelling and non-dwelling fire types had total prediction accuracies of approximately 40%. but the dwelling type had slightly better accuracy. The forest or field fire type showed high prediction accuracy: approximately 70%. The reason for this is that this fire type is outdoors, so weather influences it significantly. In addition, miscellaneous fires, which are also outdoors, had a prediction accuracy of approximately 50%. Many sources of ignition (cigarette or lighter, flame or spark, actuated equipment, and unknown), several sources of ignition (mechanical factor, negligence, and unknown), and several types of ignition materials (food; waste; and paper, wood, or hay) showed the balanced prediction accuracies across each level. The result of the prediction-model analysis for each fire characteristic was caused by the fact that the number of values equal to 0 (level 1) was higher for variables with high total prediction accuracy. For this reason, the prediction accuracies for the high levels were worse than at the low levels. In cases in which the forecast accuracies were equivalent for each level, the total prediction accuracies were fairly low; however, these cases had high usability because a prediction for each level was available.

This study has the limitation that the data had to be accurately measured at the location and the time of the fire occurrence. Additional studies are required to produce models with higher accuracies. First, more detailed analysis is required of each weather-observation location to ensure higher prediction accuracy. Second, as the prediction accuracy may change according to the desired risk level, studies that use precise boundaries of fire-occurrence risk level are required. Third, the fire characteristics associated with the weather conditions should be used for prediction, along with other variables (e.g., socioeconomic variables). Some researchers have commented on the necessity of mixed and multidimensional studies of complex interactions among people, the environment, and protective policies[2,17]. Until this study, weather conditions not been significantly considered in urban fire prediction. In the future, prediction models from other aspects can be established by analyzing factor such as weather information.

References

1. L. Yang, C. W. W. Dawson, M. R. R. Brown, and M. Gell, “Neural Network and GA Approaches for Dwelling Fire Occurrence Prediction”, Knowledge-Based Systems, Vol. 19, pp. 213-219, 2006.
2. J. Corcoran, G. Higgs, D. Rohde, P. Chhetri, J. Corcoran, G. Higgs, D. Rohde, and P. Chhetri, “Investigating the Association between Weather Conditions, Calendar Events and Socio-Economic Patterns with Trends in Fire Incidence: An Australian Case Study”, Journal of Geographical Systems, Vol. 13, pp. 193-226, 2011.
3. M. Holmes, Y. Wang, and I. Ziedins, “The Application of Data Mining Tools and Statistical Techniques to Identify Patterns and Changes in Fire Events”, 2009.
4. Z. Wang, X. Zhang, and B. Xu, “Spatio-Temporal Features of China’s Urban Fires: An Investigation with Reference to Gross Domestic Product and Humidity”, Sustainability, Vol. 7, pp. 9734-9752, 2015.
5. D.-W. Song, K.-S. Kim, and S.-K. Lee, “Statistical Analysis of the Relation between Fire Occurrence and Weather Factors”, ISSN International Journal of Engineering and Technology, Vol. 9, pp. 2319-8613, 2017.
6. K. D. Kotnour, “Statistics: The Exploration and Analysis of Data”, Technometrics, Vol. 30, pp. 460-461, 1988.
7. H. J. Kim, and J. S. Yu, “On a Method for Computing Quantiles of Grouped Data”, Journal of The Korean Data Analysis Society, Vol. 10, pp. 3453-3464, 2008.

8. P. Cortez, and A. Morais, “A Data Mining Approach to Predict Forest Fires using Meteorological Data”, Proc. 13th Port. Conf. Artif. Intell., pp. 512-523, 2007.
9. J.-S. Chou, and A.-D. Pham, “Enhanced Artificial Intelligence for Ensemble Approach to Predicting High Performance Concrete Compressive Strength”, Construction and Building Materials, Vol. 49, pp. 554-563, 2013.
10. X.-H. Meng, Y.-X. Huang, D.-P. Rao, Q. Zhang, and Q. Liu, “Comparison of Three Data Mining Models for Predicting Diabetes or Prediabetes by Risk Factors”, The Kaohsiung Journal of Medical Sciences, Vol. 29, pp. 93-99, 2013.
11. D. Delen, G. Walker, and A. Kadam, “Predicting Breast Cancer Survivability: A Comparison of Three Data Mining Methods”, Artificial Intelligence in Medicine, Vol. 34, pp. 113-127, 2005.
12. G. K. F. Tso, and K. K. W. Yau, “Predicting Electricity Energy Consumption: A Comparison of Regression Analysis, Decision Tree and Neural Networks”, Energy, Vol. 32, pp. 1761-1768, 2007.
13. L.-J. Wang, M. Guo, K. Sawada, J. Lin, and J. Zhang, “A Comparative Study of Landslide Susceptibility Maps Using Logistic Regression, Frequency Ratio, Decision Tree, Weights of Evidence and Artificial Neural Network”, Geosciences Journal, Vol. 20, pp. 117-136, 2016.
14. T. Guo, and J. V Milanovic, “Probabilistic Framework for Assessing the Accuracy of Data Mining Tool for Online Prediction of Transient Stability”, IEEE Transactions on Power Systems, Vol. 29, pp. 377-385, 2014.
15. M. Jajroudi, T. Baniasadi, L. Kamkar, F. Arbabi, M. Sanei, and M. Ahmadzade, “Prediction of Survival in Thyroid Cancer Using Data Mining Technique”, p. 13, 2014.
16. Y. Zhu, C. Xie, B. Sun, G.-J. Wang, and X.-G. Yan, “Predicting China’s SME Credit Risk in Supply Chain Financing by Logistic Regression, Artificial Neural Network and Hybrid Models”, Sustainability, Vol. 8, p. 433, 2016.
17. C. R. Jennings, “Social and Economic Characteristics as Determinants of Residential Fire Risk in Urban Neighborhoods: A Review of the Literature”, Fire Safety Journal, Vol. 62, pp. 13-19, 2013.