Article

A Deep Learning Model Development to Predict Safety Accidents for Sustainable Construction: A Case Study of Fall Accidents in South Korea

Ji-Myong Kim ¹ , Kwang-Kyun Lim ², Sang-Guk Yum ³ and Seunghyun Son ^{1,*}

¹ Department of Architectural Engineering, Mokpo National University, 1666 Yeongsan-Ro, Cheonggye-myeon, Muan-gun 58554, Korea; jimy@mokpo.ac.kr

² Department of Railroad Management, Songwon University, Gwangju 61756, Korea; lim.kwangkyun@gmail.com

³ Department of Civil Engineering, Gangneung-Wonju National University, Gangneung 25457, Korea; sy2509@columbia.edu

* Correspondence: seunghyun@khu.ac.kr; Tel.: +82-61-450-2457

Abstract: So far, studies for predicting construction safety accidents have mostly been conducted by statistical analysis methods that assume linear models, such as regression and time series analysis. However, it is difficult for this statistical analysis method to reflect the nonlinear characteristics of construction safety accidents determined by complex influencing factors. In general, deep learning techniques are used to analyze the nonlinear characteristics of complex influencing factors. Therefore, the purpose of this study is to propose a framework for developing a deep learning model for predicting safety accidents for sustainable construction. For this study, 1766 cases of actual accidents were collected by the Korea Occupational Safety Authority (KOSHA) over the 10-year period from 2010 to 2019. Eight factors influencing accident prediction such as medical day, progress rate, and construction scale were selected. Subsequently, the predictive power between deep learning models and conventional multi-regression models was compared using actual accident data at construction sites. As a result, a deep neural network (DNN) improved predictive power by 9.3% in mean absolute error (MAE) and 10.6% in root mean square error (RMSE) compared to a conventional multi-regression model. The results of this study provide guidelines for the introduction of deep learning technology to construction safety management.

Keywords: deep learning algorithm; deep neural network; artificial intelligence; construction safety accident; safety management; construction site

Citation: Kim, J.-M.; Lim, K.-K.; Yum, S.-G.; Son, S. A Deep Learning Model Development to Predict Safety Accidents for Sustainable Construction: A Case Study of Fall Accidents in South Korea. *Sustainability* **2022**, *14*, 1583. <https://doi.org/10.3390/su14031583>

Academic Editors: Amir Mosavi and Ripon Kumar Chakraborty

Received: 23 December 2021

Accepted: 28 January 2022

Published: 29 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the 4th Industrial Revolution, technology known as big data statistical analysis, artificial intelligence, robot engineering, the Internet of Things, 3D printing, unmanned transportation, nanotechnology, and cutting-edge convergence technology are being widely applied in a variety of fields [1]. Accordingly, the construction industry is also striving to shift to a new paradigm that involves periodically photographing a construction site with drones to track the site process and worker status in real time, or to try to create a digital twin-based construction site. The adoption of new technologies in the construction industry is slower than it is in other industries such as healthcare, education, finance, etc. [2]. As such, although state-of-the-art safety technology is being applied to construction sites more often, albeit partially, the accident mortality rate in the construction industry remains higher than that of other industries, and it is still generally classified as a dangerous industry [1].

In addition, although the government's continuous reduction of construction site accidents through various policies is raising the level of safety management, it has yet to achieve the desired results due to the lack of safety consciousness culture among companies

and construction workers [3]. So far, studies for predicting construction safety accidents have mostly been conducted by statistical analysis methods that assume linear models, such as regression and time series analysis.

For example, Ryu et al. [4] analyzed the actual damage record of construction sites in the loss prediction model for building construction projects using insurance claim payments. This identified the risk factors for damage. Kim et al. [5] conducted multiple regression analyses on bridge construction in a loss assessment for sustainable industrial infrastructure.

Kim et al. [6] derived economic loss indicators from quantifying the third-party loss in building construction sites utilizing claims payouts collected from major domestic insurers. Ahn et al. [7] analyzed the record of third-party damage and subsequent financial losses in bridge construction management in their paper to identify objective and loss-related risk indicators. These studies adopted statistical analysis methods. As a result, the adjusted R² was found to be 0.315 to 0.471. In other words, parts whose patterns have not been identified can be explained by hidden indicators, which requires a more reliable analysis method.

These statistical analysis methods cannot reflect the non-linear characteristic of construction safety accidents determined by complex influence factors. In the past, many studies have focused on experience, intuition, and theoretical approaches due to difficulties in securing safety accident data and limitations in analysis methods. Recently, however, data accumulation and information disclosure such as safety accident data have been carried out. Therefore, the object of this study is deep learning model development to predict safety accidents for sustainable construction. This study seeks to use deep learning techniques to quantify construction safety accident risk and propose ways to reduce safety accident risk in the construction industry.

Furthermore, we review the framework of construction safety accident prediction using deep learning techniques. In addition, we examine the effectiveness of the introduction of deep learning technology in construction safety accident prediction by comparing it to the existing construction safety accident prediction method.

This study is conducted as shown in Figure 1.

Figure 1. Structure of research.

First, safety accident-related variables to be applied to this study are investigated through a literature review. In addition, the necessity of this study is presented by considering the problems of the existing method to predict the risk of safety accidents.

Second, the purpose of this study is clearly presented based on the previously investigated contents.

Third, the method of this study is clearly explained. Data collection and input variables to be applied to the model are described in detail.

Fourth, a deep learning algorithm model is established. In this case, the optimal combination is derived through changes in hyper-parameters.

Fifth, the results and discussions of this study are clearly described.

Finally, the implications of this study and further studies in the future are described in the conclusion.

2. Literature Review

The purpose of this study is to develop a safety accident prediction model at a construction site based on a deep learning algorithm. To this end, this section conducts a literature review. Details consist of (1) safety accumulations among South Korean construction workers, (2) reviews in the construction safety literature, (3) disadvantages of externalizing methods of analysis, and (4) advances in learning.

2.1. Safety Accidents among South Korean Construction Workers

According to the Korea Occupational Safety and Health Agency's 2018 announcement, the total number of industrial accidents in the construction industry in 2018 was 6486 (29.16%), the second largest number among all industries, but it had the highest mortality rate, at 26.61% [8]. The number of casualties in the construction industry increased by 7.2% compared to the previous year, and the main causes of accidents were pinching, collapsing, bumping, and trampling, cumulatively accounting for 32% of accidents. Furthermore, these statistical data are data that have not been properly reported, as many safety accidents in this industry are concealed [3].

For example, the industrial accident rate in South Korea in 2017 was 0.84%, which was relatively low compared to the values of 3.46% in the UK and 3.88% in the US. However, comparing the number of deaths, Korea was highest among these three countries at 1.9‰, much higher than the corresponding values for the UK (0.16‰) and the US (0.95‰) [8]. These numbers are conflicting statistics according to Heinrich's Accident Triangle, which shows the relationship between the number of accidents that lead to serious, minor, and unconceived accidents, indicating that many minor and underived accident risks are not reported. Thus, accurate data collection and analysis of potential risks are essential aspects of preventing safety accidents at construction sites. In addition, the construction industry is classified as high risk compared to other domestic industries. Specifically, the construction industry has higher accident and death rates of 0.36% and 0.85‰, respectively, than other domestic industries. One of the reasons for this is that many high-rise buildings, which have high rates of death and permanent disability, have recently been built. Furthermore, construction work, which is dominated by outside work, naturally involves no protection from weather conditions such as heat, rain, and strong winds. In addition, the risk of accidents is large due to the high proportion of work occurring on roofs, ladders, and scaffolding, as well as lifting work, all of which have a high risk of safety accidents, such as fall and collapse [9].

2.2. Reviews in the Construction Safety Literature

There have been a number of studies aiming to prevent construction safety accidents occurring at construction sites; key cases include research on safety accident prevention measures [10–12], a major accident cause analysis [13], an accident case study [14], and safety accident occurrence structure analyses [15,16]. These past works have conducted post-analytical studies targeting specific types of accidents, then suggested prevention and countermeasures based on the results. These studies cannot be applied to ongoing activities to prevent safety accidents and minimize damage in advance while considering the conditions of construction sites, wherein dynamic and diverse companies work. Nevertheless, due to the nature of the construction industry, such as the limited budget for safety management at construction sites, the involvement of manpower, and the large number of external works, efforts to prevent and reduce risk in advance are essential.

Therefore, it is necessary to accurately calculate potential safety accidents in advance. Such accident estimation will help prevent and reduce damage by identifying the types of safety accidents that may occur and the risks that should be managed first at each

construction site. This will ultimately contribute to increased profits for construction projects and further reductions of safety accidents in the construction industry as a whole. Consequently, this study intends to examine the feasibility of applying deep learning techniques to predict potential construction safety accidents at construction sites in advance.

2.3. Disadvantages of Existing Methods of Analysis

Most prior studies on safety accident analysis and safety accident risk prediction at construction sites have been based on traditional statistical methods such as regression analysis, fuzzy theory, literature reviews, expert opinions, and expert interviews. Specifically, safety accidents in construction projects have been examined through expert consultations, interviews, and expert surveys such as AHP (analytic hierarchy process), Delphi survey, etc. due to the difficulty of securing data [17–21]. However, this research method requires substantial amounts of time and money, and accurate communication is impossible in some cases due to difficulties that some people may have with understanding or answering questions. It is also difficult to resolve any conflicting ideas among respondents. Regression analysis is mainly used for data-based research.

To solve the problem of these existing statistical analysis methods, machine learning techniques, which are methods by which machines learn by themselves from the data provided, must be utilized. In particular, among machine learning techniques, deep learning algorithms are a way of learning data by utilizing information input/output layers similar to neurons in the brain. These deep learning algorithms classify datasets and find correlations among themselves to make predictions more accurately. Furthermore, the deep learning algorithm can have a more complex structure than other artificial intelligence techniques, such as machine learning and fuzzy logic, so its availability can be increased when complex and large amounts of data are secured in the future [22]. In addition, considering the irregularity of safety accident data, a deep learning algorithm that can have a more multifaceted structure can explain safety accident data better than other artificial intelligence techniques [23]. Therefore, this study intends to present a safety accident analysis methodology using a deep learning algorithm.

2.4. Advantages of Deep Learning Techniques

These studies define the correlation between safety accidents and various factors, i.e., project information, natural hazards, geographical vulnerability, building environment, and socio-economic factors, and estimate the severity and frequency of safety accidents; the risk of safety accidents can be calculated through this process [4,5,7,24–27]. However, due to the development of ICT technology, various materials that have not been used often in the past are being produced and accumulated in large quantities at construction sites. In addition, the recent geometric development of computer software and hardware has made it possible to analyze enormous amounts of information in a short time. Based on these advances in data acquisition and analysis technology, big data analysis technologies such as machine learning and deep learning have begun to attract attention as representative and new approaches [6].

These machine learning and deep learning techniques are advantageous for analyzing factors related to safety accidents using large amounts of data, and are already being actively applied in other fields. Despite these advantages, the utilization of such techniques remains low in the field of construction safety [28]. In this field, these techniques can make it possible to estimate the relationship with risk factors related to safety accidents that are not defined in the existing statistical models by using heterogeneous and large amounts of data. The risk indicators were identified through literature reviews and statistical analyses of many studies, but the predictive power of the model is still between 0.2 and 0.5 [17–21]. In other words, this model can explain about 20–50% of the relationship with safety accidents through the prescribed indicators, whereas the remaining 50–80% of the relationship is explained by unidentified indicators. Thus, it is necessary to advance the estimation of safety accidents by identifying these hidden indicators.

It can also be said that machine learning and deep learning techniques in a big data environment are likely to develop further. In particular, when combined with big data such as safety sensor information, weather information, the Internet of Things, and unmanned aerial vehicles, which have recently come to be widely used in construction sites, not only can the accuracy of safety accident calculation be further improved, but it is also ultimately expected that the risk of safety accidents at construction sites can be reduced [29–31].

3. Research Objectives

As previously explained, existing safety accident analysis methods use statistical methods. However, these statistical approaches do not reflect a variety of influence factors because they are assumed to be a linear model. In particular, multiple regression analyses can be easily analyzed with fewer data, but when the variability (or error) of the data is large, there is a problem of generating predictions that differ greatly from the actual value. Machine learning, which has recently attracted attention in the field of engineering, is an area of artificial intelligence technology, and active research is underway in several fields, showing significant results. Furthermore, the limitations of existing statistical analysis models can be supplemented in that machine learning is a nonlinear estimation technique.

The ultimate aim of this study is to provide a framework for developing a construction site safety accident prediction model using a deep learning algorithm. The aim of this study is to develop a construction site safety accident prediction model using deep learning algorithms based on the safety accident data occurring at the construction site. The detailed objectives are to (1) explore the safety accidents that occur at real construction sites in South Korea and create a construction site safety accident prediction model using deep learning algorithms as a case study, and (2) compare the models by comparing the prediction error of the deep learning algorithm model and the model using existing statistical techniques to verify the validity of the introduced model.

4. Methodology and Data Collection

The detailed research steps are shown in Figure 2. First, based on past studies, this study investigated indicators and models related to safety accidents. Second, safety accidents that occurred at construction sites were collected as input and output variables.

Figure 2. Methodology.

The deep learning model uses Python 3.8 and the statistical model analysis tool was made using IBM Statistical Package for the Social Sciences (SPSS) V23. Third, each model (the deep learning model and the statistical model) was developed. Finally, the MAE (mean absolute error) and the RMSE (root mean squared error) were calculated to compare the models.

This study used safety accident data from 2010 to 2019 from the Korea Occupational Safety and Health Agency (KOSHA) to secure data on safety accidents that occurred at construction sites in Korea. KOSHA is a government agency that was founded with the Korea Occupational Safety and Health Agency Act in 1987. This institution was established to help employers prevent accidents and ensure that workers work in a safe and healthy manner. Representative responsibilities include research and development related to safety, education on occupational safety and health, and the dissemination of industrial accident

prevention technologies. To this end, safety accidents occurring in various industrial sites are counted and converted into data.

The collected data include the construction scale, number of employees, worker nationality, accident type, construction progress at the time of the accident, accident occurrence date, accident occurrence time, accident day, employment type, site address, accident summary, and medical treatment day. The dependent variable was the number of days of treatment for the accident. To unify the types of accidents and quantify the severity of each accident, among the data, death was excluded, and the type of accident was limited to accidents caused by falling. The collected data did not contain any personal information from the accident.

In this study, eight variables were selected in reference to the previous study, as shown in Table 1. Kim et al. [3] selected the days of treatment, progress rate, construction scale, and employment types as variables in their paper. Kim et al. [26] selected the progress rate, construction scale, and worker's carelessness as variables in their paper. In particular, Kim et al. [3] selected days of treatment as a dependent variable to analyze the risk of accidents in migrant workers, and independent variables such as progress rate, construction scale, classification of occupations, day of the week, accident time, and employment type. As a result, the independent variable was statistically significant at level 0.01.

Table 1. Descriptions of the variables.

	Variable	Type	Description
Input	Progress rate (%)	Numerical	Process rate of construction site at the time of an accident
	Construction scale	Nominal	Small, medium, and large construction sites, according to the standards of the Korea Occupational Safety and Health Act
	Number of employees	Numerical	Number of employees employed by the company
	Status of employment	Nominal	Regular workers and non-regular workers
	Nationality of workers	Nominal	Non-migrant and migrant workers
	Day of week	Nominal	Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, and Sunday
	Accident time	Nominal	Dawn, morning, afternoon, and evening and night
	Occupational classification	Nominal	Occupation classification according to Korean Standard Statistical Classification
Output	Medical day	Numerical	Number of treatment days

According to Kim et al. [3], the risk increases as the progress rate and construction scale increase. In addition, they said that the risk of accidents may increase depending on the classification of occupations, day of the week, accident time, and status of employment. As such, their study found that factors such as progress rate, construction scale, number of employees, status of employment, nationality of workers, day of the week, accident time, and classification of occupations influence the risk of accidents at construction sites.

Therefore, the model in this study was applied with eight variables. In the future, it is intended to increase the reliability of the model by reflecting factors such as the difficulty of the construction, worker safety awareness, organizational culture, etc.

In total, eight indicators were used in the safety accident prediction model, and descriptions of these variables are listed in Table 1.

As the process rate increases in construction, the risk of safety accidents increases due to the increased construction complexity caused by collisions between various processes [25,32]. Therefore, this rate has a close correlation with safety accidents in a construction project. This study asserts that the degree of risk is reflected using the construction site process rate at the time of the safety accident.

The scale of construction work and the number of employees have statistically significant relationships with safety accidents, and these are widely used as indicators of the

risk of safety accidents [24,33,34]. This study adopted the scale of construction work being divided into two units as a nominal variable, in accordance with the Korean Occupational Safety and Health Act. Specifically, each site was classified according to the total construction amount, i.e., (1) a small construction site with a total construction cost of less than KRW 2 billion, (2) a medium-sized construction site with a total construction cost of KRW 2 billion to 12 billion, and (3) a large construction site with a total construction cost of more than KRW 12 billion. The number of employees maintained by the company was also entered.

The employment type was divided into non-regular workers and regular workers. Research has suggested that non-regular construction workers are more susceptible to safety accidents than regular construction workers [3]. Therefore, employment status is an important indicator for the risk of safety accidents. This study defines the indicator as a nominal variable, i.e., (0) regular worker and (1) non-regular worker.

Moreover, the nationality of the worker is an important indicator of the vulnerability of workers to safety accidents. This is because migrant workers have relatively shorter careers than non-migrant workers, and they are often more exposed to safety accidents as they work in small and medium-sized workplaces [35–37]. This study adopted the indicator as a nominal variable, i.e., (0) non-migrant worker and (1) migrant worker.

Further, the day and time are important indicators that indicate the risk of a safety accident [38,39]. The work at a construction site is characterized by changes in the work content and intensity depending on the day and time of the week, because construction workers tend to move fluidly according to the working environment of the construction site [3].

For example, safety accidents frequently occur in the middle of the morning and early afternoon, when work is intensive. This is also because the accident rate increases during times of heavy workload or times of high difficulty. Consequently, in this study, time was divided into nominal variables, i.e., (1) dawn (00:00~06:00), (2) evening and night (18:00~24:00), (3) afternoon (13:00~18:00), and (4) morning (06:00~12:00). The day of the week was also used as a nominal variable, classified as (1) Monday, (2) Tuesday, (3) Sunday, (4) Wednesday, (5) Friday, (6) Saturday, and (7) Thursday.

Occupational classification is used to evaluate the risk of work according to occupational classification [3]. The Korean Standard Statistical Classification was used for occupational classification, which characterizes risk factors according to the job and provides comprehensive management guidelines for similar jobs [40]. Therefore, it was used as a factor indicating the risk of safety accidents because it reflects the difficulty and intensity of work according to the work classification.

The occupational classification was divided into nominal variables, i.e., (1) equipment, machine operating, and assembly workers; (2) professionals and related workers; (3) craft and related trade workers; (4) managers; and (5) elementary workers. The descriptive statistics of the variables are listed in Table 2.

This study used the data on safety accidents from 2010 to 2019 from the Korea Occupational Safety and Health Agency (KOSHA). We continue to strive to secure actual safety accident data at construction sites. In the future, additional data will be obtained continuously to increase the reliability of the model.

Table 2. Descriptive statistics of variables.

Variable	N	Min.	Max.	Mean	Std. Deviation
Input					
Progress rate (%)	1766	1.00	100.00	51.93	26.60
Construction scale (KRW)	1766	1.00	3.00	1.34	0.62
Number of employees	1766	1.00	2000.00	25.83	102.75
Status of employment	1766	0.00	1.00	0.89	0.31
Nationality	1766	0.00	1.00	0.05	0.23
Day of week	1766	1.00	7.00	4.02	2.05
Accident time	1766	1.00	4.00	3.45	0.60
Occupational classification	1766	1.00	5.00	3.78	0.99
Output					
Medical day	1766	1.00	361.00	95.28	37.25

5. Construction of Deep Learning Algorithm Model

In this study, a design process for deriving optimal input variables and parameters was constructed to develop a deep learning-based accident risk prediction model. The predictive model design process is shown in Figure 3.

Figure 3. Overview of the predictive model design process.

As shown in Figure 3, when the input variable derivation process is performed and the optimal parameter derivation process is completed, a deep learning model based on the deep neural network is developed. The final developed model is verified and evaluated for effectiveness. The first step in the process of deriving optimal input variables is to identify important variables from the actual data, and for this purpose, prior research and expert knowledge were used. In the process of deriving optimal parameters, deep learning algorithms using deep neural networks were used.

First, it is necessary to divide actual data into learning and verification data, and to set the range of parameters. In order to optimize the DNN model, the number of inputs, hidden layers, number of nodes in the hidden layer, dropout, activation function, optimizer, batch, epochs, etc. should be set directly to determine the optimal combination for the model. MAE and RMSE are used to evaluate the predictive performance of the developed model. Finally, when the optimal DNN model is determined, the predictive performance is evaluated compared to the existing method, MRA.

MAE and RMSE are two evaluation indicators that express errors through comparisons between predicted values and actual values using artificial neural networks [41]. MAE is obtained by taking the difference between the predicted value and the actual value as

an absolute value and averaging it, and the closer the MAE value is to 0, the smaller the error, so it can be said that the predictive power of the model is high. Meanwhile, RMSE is a measure commonly used to deal with the difference between the predicted value of the model and the value observed in the real environment. RMSE is used to synthesize residuals into one measure. The closer the RMSE value is to 0, the smaller the error, again indicating that the predictive power of the model is high.

In this study, we focused on the development of models using deep learning algorithms. No separate predictive power was produced in this study. MAE and RMSE are commonly used as model evaluation metrics in machine learning and deep learning. The lower the MAE and RMSE, indicators of the difference between predicted and real values, respectively, the higher the predictive power of the model. This study adopted Python 3.8 for model development, and it also used the Keras and Scikit-Learn libraries.

5.1. Deep Learning Algorithm

As AI continues to evolve, its value is continually increasing with the emergence of new algorithms. Neural networks are one of the deep learning algorithms. Neural networks were introduced as algorithms that mimic the neural system, and they solve problems using a layer of connected nodes. The multilayer perceptron (MLP) is mostly used for regression problems as well as classification and recognition [42]. This study used a deep learning technique, which is widely used in the field of recognition and prediction. A deep learning model has basically the same structure as an artificial neural network, but it has the advantage of having two or more hidden layers, so it can be applied to various type of data [43].

Deep learning is one of the techniques of machine learning, and it refers to a method of performing type classification or regression of input data. In general, deep learning has a neural network composed of several layers, and it has various structures [44].

Deep learning involves various models, such as the deep neural network (DNN), convolutional neural network (CNN), recurrent neural network (RNN), autoencoder (AE), and generative adversarial network (GAN), depending on the processing method and structure. DNNs are usually standard neural networks with depth. There are hidden layers between the input and output layers, and the number of hidden layers determines the depth of the neural network. There is no specific number that a neural network classifies as deep, but in general, if the credit assignment path depth exceeds 2, it is regarded as deep [45].

DNNs are trained to model complex nonlinear relationships, since each layer of the multiple layers is designed to identify a specific function [46]. DNNs have the advantage of being capable of various types of modeling, like a general artificial neural network. However, they also have the disadvantage of being vulnerable to overfitting. Consequently, they are generally utilized for prediction and classification in various fields [47]. This study developed a safety accident prediction model at a construction site using DNN in consideration of the input data format and output format as well as various modeling possibilities.

5.2. Data Preprocessing

The total number of data collected in this study was 1766, and eight input variables were used: progress rate, construction scale, number of employees, status of employment, worker nationality, day of week, accident time, and occupational classification. In this study, data were preprocessed by scaling the data using the z-score normalization method, which is a method that normalizes data using the mean and standard deviation. Before data pre-processing, data duplication, invalidity, and duplication were investigated, and it was confirmed that there were no abnormalities. Six nominal variables, including construction scale, status of employment, worker nationality, day of the week, incident time, and occupational classification, were entered as a single variable because they did not require normalization. Two numerical variables, the progress rate and the number of employees, were normalized. The z-score normalization method was adopted in this study. This was done to

avoid the problem of data corresponding to outliers. In this work, z-score normalization was applied to reflect a wide variety of cases of construction safety accident data [48].

Data preprocessing is the process of adjusting the range of data for easy comparison and applying units of quantities and resources that are difficult to compare. The input data preprocessed through the normalization process were divided into three types (i.e., learning, verification, and testing). Training data (learning and verification data) are data for a deep learning algorithm model to learn based on input data. In order to check whether the performance of the trained model is optimal, an optimal deep learning model is constructed and subjected to a verification process with verification data; the verification data are included in the training data. It is necessary to evaluate whether the verified optimal learning model is suitable for predicting the problem. At this time, the predicted result of the model and the result of the test data are compared to calculate and evaluate the error rate and fitness, and the data used for this step are the test data. The data used in this study were divided into training data (70%) and test data (30%); in addition, 30% of the training data were used as validation data. All data were chosen randomly.

In the model of this study, the data were scaled using the z-score normalization method to preprocess the data. If noise data are generated during the preprocessing process, the process may be performed by interval estimate or by binning through a smoothing method. In addition, if a new variable occurs in predicting safety accidents in the future, the DNN model of this study can be transferred learning and recalibrated to suit the problem to be newly solved.

5.3. Network Structure Scenario and Hyper-Parameter Tuning

The DNN model optimizes the model by updating the weights of each neural network node through a backpropagation algorithm. In order to optimize the DNN model, the number of inputs, hidden layers, number of nodes in the hidden layer, dropout, activation function, optimizer, batch, epochs, etc. should be set directly to determine the optimal combination for the model [49]. Since the optimal combination value differs according to the input/output variables of each model, the optimal value of the model is derived through a trial-and-error method. The dropout is a normalization penalty to prevent overfitting.

In a deep learning algorithm, when a model has many input variables or when a learning model becomes complicated due to a large number of hidden layers and nodes, an overfitting phenomenon may occur. Overfitting refers to a phenomenon in which the performance of a deep learning model is degraded due to an increase in errors in the actual data or the test data caused by excessive learning of the training data. Dropouts can be used to prevent this [50].

The activation function is a method of finding the weight that minimizes the cost function of each node. The optimizer is a way to make the learning both stable and fast. The batch is responsible for grouping the data into sets for efficient computation by large neural networks. The epoch defines the number of transcriptions [47,49]. In this study, considering the number of learning data and related past studies, three hidden layers as well as 0 or 0.2 for dropout were determined [51–53]. The results of learning are shown in Table 3.

The phenomenon of overfitting a model to the learning data is a major issue that reduces the model's performance. Overfitting reduces the predictive power of the model due to the overlearning of learning data. To prevent this, dropout rates should be adjusted. It was determined that dropout rates do not make much difference by comparing the results of the training data, the results of the validation data, and the results of the test data, as shown in Table 3.

Table 3. Learning result.

Network Structure Scenario	Dropout (0)		Dropout (0.2)	
	MAE	RMSE	MAE	RMSE
5-5-5	26.21	35.18	26.17	34.63
10-10-10	26.54	35.14	26.37	34.84
25-25-25	26.87	35.11	26.56	35.04
37-37-37	26.76	35.20	26.48	35.01
50-50-50	26.66	35.29	26.40	34.98
75-75-75	26.70	35.22	26.39	34.88
100-100-100	26.73	35.16	26.37	34.78
151-151-151	26.64	35.19	26.45	34.92
200-200-200	26.55	35.22	26.53	35.07
251-251-251	26.53	35.21	26.57	35.11
300-300-300	26.52	35.21	26.62	35.15
351-351-351	26.33	35.08	26.68	35.22
400-400-400	26.15	34.96	26.75	35.30
436-436-436	26.32	34.91	26.63	35.23
451-451-451	26.33	34.93	26.64	35.22
500-500-500	26.51	34.90	26.52	35.14

The model with the minimum MAE and RMSE values for each dropout and node scenario was chosen as the final model. The learning results are presented in Table 3. As the number of nodes in the hidden layer increased, the MAE and the RMSE tended to increase gradually, and when the number of nodes was 400-400-400, the MAE and the RMSE were both found to be minimal. Consequently, for the network structure, dropout was determined as 0 and the number of nodes was determined as 400-400-400. The batch was 5, the activation function was the ReLU (rectified linear unit) function, and the optimizer used the Adaptive Moment Estimation (Adam) method. The epoch was set to 1000 times, but the learning was stopped where there was no further improvement in the loss value. As the activation function, the ReLU function is a method designed to solve the vanishing gradient problem, which is a disadvantage of the existing sigmoid function. The ReLU function returns 0 if the input value is less than 0, and it sends the input value as it is if it is greater than 0 [46,51]. The Adam method was used as the optimization algorithm. This optimizer is a method that was proposed in 2015, and it is a commonly used algorithm due to its advantages, such as its ease of implementation, efficiency of calculation, and excellent performance. This algorithm is a first-order gradient algorithm based on a stochastic objective function by adding the concept of moment [52].

In this study, the ReLU activation function was used in the hidden layer, and the activation function for the output layer used the softmax function. The activation function is used to normalize and transfer input values between 0 and 1 when they pass through one layer and transfer to the next layer. Particularly, the ReLU activation function does not cause gradient vanishing that occurs in the sigmoid function, and the slope value is 1 as long as the input value is greater than 0, so learning in backpropagation is fast. For this reason, it is the most commonly used activation function in recent years. Since the data collected in this study did not have a value less than 0, the fast-learning ReLU activation function was used in the hidden layer.

In addition, the main reason for using the softmax function in the output layer in this study was to interpret the result value as a probability value. The biggest advantage of the softmax function is that the sum of probabilities is 1, so it is easy to recognize which classification has the highest probability. Due to these characteristics, the use of softmax can obtain the effect of normalization, so it is an activation function mainly used in the output layer.

Dropout was used on the input and hidden layers, not on the output layers. The learning rate was set at 0.1 for first 30 epochs, and 0.01 for next 70 epochs. It was set to 0.001 until the end. Epochs are the number of times the model is repeatedly learned. In

this study, we trained it by changing the number of epochs, and learning became the most efficient when the number of epochs was specified as 1000.

5.4. Validation of the Model

The final model was decided based on the determined network structure and hyper-parameters. In this study, the Adam (Adaptive Moment Estimation) method was used as an optimizer. The Adam method can be implemented simply and has the advantage of having few memory requirements due to its efficient operation.

The final model is detailed in Table 4. To verify the model, the predicted value and actual value (medical day) were calculated using the verification data and test data in the final DNN model, and the MAE and the RMSE were compared. In addition, using multiple regression analysis (MRA), which is an existing calculation method, the MAE and the RMSE were calculated and compared with the same data as the final DNN model. The MRA model was built using IBM SPSS Statistics 23 for Windows.

Table 4. Determined network structure and hyper-parameter.

Category	Composition	Value
Network structure	Layer	3
	Node	400-400-400
Hyper-parameter	Optimizer	Adam (Adaptive Moment Estimation) method
	Activation function	ReLU (rectified linear unit) function
	Dropout	0
	Batch Size	5
	Epoch	1000

The results of the analysis of the evaluation indicators for each model are presented in Table 5. The results of the validation data of the DNN model (MAE: 26.05, RMSE: 35.01) showed no significant difference compared to the training data results (MAE: 26.15, RMSE: 34.96); this indicates that the problem of overfitting is insignificant in this model. The results of the MRA and DNN models were MAE 29.98 and RMSE 39.79 for MRA, and MAE 27.42 and RMSE 35.99 for DNN. This indicates that the prediction error of the MRA model was larger than that of the DNN model, with MAE at 9.3% and RMSE at 10.6%; that is, the DNN model also had a lower prediction error rate than MRA in MAE by 9.3% and in RMSE by 10.6%.

Table 5. Comparison of analysis results of verification data and test data.

Model	Validation		Test	
	MAE	RMSE	MAE	RMSE
DNN	26.05	35.01	27.42	35.99
MRA	-	-	29.98	39.79
DNN/MRA (%)			−9.3%	−10.6%

6. Results and Discussion

As discussed earlier, meaningful results were obtained through the process of optimizing the network structure and hyper-parameters by borrowing the DNN algorithm, which is one of the deep learning algorithms. As a result of the model's prediction, the DNN model showed that the MAE was 27.42 and the RMSE was 35.99. The MRA model showed a MAE of 29.98 and an RMSE of 39.79. The analyzed results show that the DNN model improved its predictive power by 9.3% in MAE and 10.6% in RMSE over the MRA model. The reason for these results is that the existing MRA model is a statistical analysis

method that assumes a linear model, so the non-linear characteristics of the data due to complex influencing factors cannot be reflected. Contrary to this, the DNN model can reflect non-linear characteristics that are difficult for humans to check by learning a large amount of data by a machine. Therefore, it was judged that the DNN model was more suitable than the MRA model for adequately reflecting the nonlinear characteristics of the construction site safety accident data.

Deep learning techniques will allow us to predict safety accidents in construction projects more effectively than the MRA model. Using the methodology presented in this study, an optimal deep learning prediction model for a user's desired problem can be constructed. Hence, the ultimate goal of this study, the framework proposal for the development of a deep learning model for the prediction of construction safety accidents, presented significant results. In addition, using this deep learning model, other users will be able to predict safety accidents at construction sites. For instance, in a construction site, the risk of an accident can be predicted through the DNN model, and management for the prevention of accidents can be advanced. In addition, the ordering party can predict accident risk in advance as part of the safety measure and use it to prepare a reserve fund, safety management budget, and annual budget for accident response. Furthermore, when insurance is actively considered as an accident risk transfer strategy, it can be used as a guideline for the selection of an appropriate premium. Moreover, since this model can be used to increase prediction reliability over that of the existing model, it will be possible to derive a more objective and accurate safety accident risk than the existing model. For instance, a project owner can set a level of probability that fits their assets and risk appetite. Through this, it is possible to prepare for risks in advance by transferring the risk through insurance for the expected risk amount and by signing a special contract for supplementary risks. In addition, it will be an opportunity to think about an appropriate rate by comparing it with the current insurance rate. Furthermore, it can contribute to improving the safety accident rate by enabling active safety management of construction sites through the accurate prediction of safety accidents. Using this model, the risk of safety accidents can be identified in advance, and damage can be minimized and avoided through appropriate investments for risk reduction. This will greatly reduce construction accidents by improving construction site safety, thereby contributing to the minimization of social costs resulting from human and material losses.

If the process of predicting accident risk proposed in this study is systematized, more effective safety management will be possible for accident prevention at construction sites. In the future, the safety management system to be developed will operate as shown in Figure 4.

As shown in Figure 4, first, the construction site is selected (Step 1) and then the construction site data are investigated (Step 2). Subsequently, the eight variables presented in this study are entered into the model (Step 3). Deep learning models operate using the entered variables (Step 4). The model measures the risk of accidents (Step 5). Subsequently, the risk of accidents at construction sites is reviewed (Step 6). If the risk of an accident is not identified correctly, the risk of an accident is reviewed again (feedback to Step 2). If the risk of an accident is identified correctly, its effectiveness is verified by application in the case (Step 7).

Nevertheless, this study lacks detailed data that can be used in the DNN model, even though, among the safety accident data at the construction site, the data from the Korea Occupational Safety and Health Agency were relatively accurate and large. In order to estimate safety accidents using deep learning algorithms, it is necessary to systematically build and manage safety accident data and geometry data as a database. In addition, further research is needed to advance the model by increasing the type of input variables and the amount of data. Furthermore, detailed personal information (age, gender, etc.) other than the variables used in this study could not be obtained from the KOSHA data. Further research is needed, including personal information that can be closely linked to safety accidents. Moreover, the data output of KOSHA collected for this study was the

medical day of the accident, and the accident type (mild, severe, etc.) was not considered. Although the severity of an accident can be inferred through the medical day, there is a limit to the sophisticated prediction of safety accidents, so additional research is needed to develop a model for each accident type by securing additional data.

Figure 4. Operating process for a construction safety management system.

Furthermore, the existing studies using deep learning algorithms to predict safety accidents were relatively insufficient compared to those in other areas, so the selection of the optimal model and model structure depended on trial and error. Further research is needed through an in-depth analysis of more diverse model structures and parameter values, as well as the introduction of other algorithms in addition to the DNN algorithm used in this study.

Additionally, in this study, a model was developed using the variables employed in previous studies. The developed model has a low prediction error as a result of verification and is evaluated as a meaningful research result. However, it is not possible to know the weight and significance of each variable due to the characteristics of the DNN model.

In deep learning, details in which input data are processed in multiple layers are stored in the weight of the layer. Learning is to find weight values in all layers of the neural network to map input data to an accurate target. However, some deep neural networks have tens of millions of parameters. In this case, it is very difficult to find the exact values of all parameters. The reason is that changing one parameter value affects all other parameters. Therefore, it is difficult to express it in a general formula.

This may act as a lethal restriction in logically explaining the process and rationale for future result derivation by providing unilateral prediction results through the DNN algorithm.

To compensate for this, Europe has recently introduced eXplainable Artificial Intelligence (XAI) and made it mandatory. Therefore, in future studies, XAI will be introduced to improve the reliability of the DNN model.

eXplainable AI is an AI technology that explains the causal relationship, decision-making, prediction, and classification of the model to help users understand the model results [51]. Through such AI technology, it will be possible to not only improve reliability by relieving anxiety about model results, but also develop understanding into data, models, and variables.

In addition, in this study, only the DNN algorithm was borrowed considering the complex nonlinearity of construction safety accidents. This can act as a limitation in model results and utilization. Therefore, cross-validation between models through modeling of other deep learning algorithms (e.g., CNN, RNN, etc.) is required in future research to increase the excellence and persuasiveness of model results. Such cross-validation will increase the reliability of the model results.

7. Conclusions

In this study, a deep learning model was developed to predict construction safety accidents. Deep learning technology can be a way to reduce safety incidents in construction sites. The model can improve the estimation technique of safety accidents at construction sites and improve the accuracy of prediction. The model proposed in this study is expected to develop into a more reliable model through continuous data accumulation and effectiveness verification.

The results of this study are as follows.

First, a deep learning-based safety accident prediction model was developed using the collected data. The model can predict the medical day by reflecting factors such as progress rate, construction scale, number of employees, status of employment, nationality of workers, day of week, accident time, and occupational classification. The developed model showed minimal MAE and RMSE when the number of nodes was 400-400-400. The MAE was 26.15 and RMSE was 34.96.

Second, the predictive power between the model in this work and the existing multi-regression model was compared. As a result, the MRA model showed that the MAE was 29.98 and the RMSE was 39.79. Furthermore, the DNN model showed the MAE as 27.42 and the RMSE as 35.99. The DNN model improved predictive power by 9.3% in MAE and 10.6% in RMSE compared to a conventional multi-regression model.

The model built through this study can improve the estimation technique of safety accidents at construction sites and improve the accuracy of prediction. Applying this model to construction sites can identify the risk of safety accidents in advance and minimize damage through appropriate investments to reduce risk. In the future, the model proposed in this study is expected to develop into a more reliable model through continuous data accumulation and effectiveness verification.

This study provides a framework for the introduction of DNN models in safety management and can be applied in other industries. Additionally, the development procedure and development model used for the deep learning algorithm model developed in this study can be applied to various types of construction sites and related research. This is expected to contribute to improving the safety accident rate in the overall construction industry. In the future, further research on the development of safety management system software will be conducted by developing the results of this study. If the proposed process of predicting the risk of accidents is systematized through this study, more effective safety management will be possible to prevent accidents at construction sites.

Author Contributions: Conceptualization, J.-M.K. and S.S.; methodology J.-M.K. and S.S.; validation, S.S.; formal analysis, K.-K.L. and S.-G.Y.; investigation, S.-G.Y.; resources, J.-M.K.; data curation, K.-K.L.; writing—original draft preparation, J.-M.K., S.S. and S.-G.Y.; writing—review and editing, J.-M.K.; visualization, S.-G.Y. and K.-K.L.; supervision, S.S.; funding acquisition, J.-M.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (NRF-2019R1F1A1058800).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Gledson, B.J.; Greenwood, D. The adoption of 4D BIM in the UK construction industry: An innovation diffusion approach. *Eng. Constr. Archit. Manag.* **2017**, *24*, 950–967. [CrossRef]
- IPA. *Transforming Infrastructure Performance*; Infrastructure and Projects Authority: London, UK, 2017.
- Kim, J.M.; Son, K.; Yum, S.G.; Ahn, S. Analyzing the risk of safety accidents: The relative risks of migrant workers in construction industry. *Sustainability* **2020**, *12*, 5430. [CrossRef]
- Ryu, H.; Son, K.; Kim, J.M. Loss prediction model for building construction projects using insurance claim payout. *J. Asian Archit. Build. Eng.* **2016**, *15*, 441–446. [CrossRef]
- Kim, J.M.; Kim, T.; Ahn, S. Loss assessment for sustainable industrial infrastructure: Focusing on bridge construction and financial losses. *Sustainability* **2020**, *12*, 5316. [CrossRef]
- Kim, J.M.; Ha, K.C.; Ahn, S.; Son, S.; Son, K. Quantifying the third-party loss in building construction sites utilizing claims payouts: A case study in South Korea. *Sustainability* **2020**, *12*, 10153. [CrossRef]
- Ahn, S.; Kim, T.; Kim, J.M. Sustainable risk assessment through the analysis of financial losses from third-party damage in bridge construction. *Sustainability* **2020**, *12*, 3435. [CrossRef]
- Korea Occupational Safety and Health Agency. 2018 Industrial Accident Status Analysis. 2018. Available online: <https://www.kosha.or.kr/english/index.do> (accessed on 18 September 2021).
- Chong, H.Y.; Low, T.S. Accidents in Malaysian construction industry: Statistical data and court cases. *Int. J. Occup. Saf. Ergon.* **2014**, *20*, 503–513. [CrossRef]
- Seo, J.M.; Kim, S.K. The study on measures for reducing safety accidents of excavator. *J. Korea Inst. Build. Constr.* **2008**, *8*, 127–133. [CrossRef]
- Kim, J.W.; Kim, Y.H.; Kim, J.H. A Study on the Methods of Reducing the Safety Accident through Pattern Analysis of Construction Disasters. In Proceedings of the 2010 Conference of The Korea Institute of Building Construction, Seoul, Korea, 3 November 2010; Volume 10, pp. 137–140.
- Kim, J.H. A study on prevention plan of accident through the analysis of cause for safety accident in the demolition work. *J. Reg. Assoc. Archit. Inst. Korea* **2012**, *14*, 271–278.
- Shin, D.P.; Son, C.B.; Lee, D.E. Association analysis of construction accident attributes causing fatalities. *J. Archit. Inst. Korea* **2012**, *28*, 87–94. [CrossRef]
- Wagawa, M.; Kim, J.G.; Cho, J.Y. Investigation of the Relation between Economic Transformation and Construction Safety Accidents in Japan. In Proceedings of the 2013 Conference of Korea Institute of Construction Engineering and Management, Seoul, Korea, 13 November 2013; pp. 309–310.
- Song, H.; Park, H.G.; Go, S.S. A study on the analysis of accident cause of form work using FTA (Fault Tree Analysis) system. *J. Archit. Inst. Korea* **2006**, *22*, 119–127.
- Lee, J.S.; Kim, M.J.; Choi, K.H. An efficient safety management through mechanism analysis on disasters of temporary facilities. *J. Archit. Inst. Korea* **2010**, *26*, 129–136.
- Choi, M.C. Problems of AHP analysis and development of a modified weight model. *Manag. Inf.* **2020**, *39*, 145–162.
- Gambatese, J.; AlOmari, K. Degrees of connectivity: Systems model for upstream risk assessment and mitigation. *Accid. Anal. Prev.* **2016**, *93*, 251–259. [CrossRef]
- Alomari, K.A.; Gambatese, J.A.; Tymvios, N. Risk perception comparison among construction safety professionals: Delphi perspective. *J. Constr. Eng. Manag.* **2018**, *144*, 04018107. [CrossRef]
- Gao, R. Managing Construction Safety in Chinese International Contractors' Overseas Projects: A Case in Vietnam. Ph.D. Thesis, The Hong Kong Polytechnic University, Hong Kong, China, January 2017.
- LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
- Zhong, B.; Pan, X.; Love, P.E.; Ding, L.; Fang, W. Deep learning and network analysis: Classifying and visualizing accident narratives in construction. *Autom. Constr.* **2020**, *113*, 103089. [CrossRef]
- Wang, Z.Z.; Chen, C. Fuzzy comprehensive Bayesian network-based safety risk assessment for metro construction projects. *Tunn. Undergr. Space Technol.* **2017**, *70*, 330–342. [CrossRef]
- Kim, J.M.; Kim, T.; Son, K.; Bae, J.; Son, S. A quantitative risk assessment development using risk indicators for predicting economic damages in construction sites of South Korea. *J. Asian Archit. Build. Eng.* **2019**, *18*, 472–478. [CrossRef]
- Kim, J.M.; Kim, T.; Bae, J.; Son, K.; Ahn, S. Analysis of plant construction accidents and loss estimation using insurance loss records. *J. Asian Archit. Build. Eng.* **2019**, *18*, 507–516. [CrossRef]
- Yum, S.G.; Bae, S.J.; Kim, J.M. Assessing the risk of natural disaster-induced losses to tunnel-construction projects using empirical financial-loss data from South Korea. *Sustainability* **2020**, *12*, 8026. [CrossRef]
- Jin, R.; Zou, Y.; Gidado, K.; Ashton, P.; Painting, N. Scientometric analysis of BIM-based research in construction engineering and management. *Eng. Constr. Archit. Manag.* **2019**, *26*, 1750–1776. [CrossRef]

28. Akinosho, T.D.; Oyedele, L.O.; Bilal, M. Deep learning in the construction industry: A review of present status and future innovations. *J. Build. Eng.* **2020**, *32*, 101827. [\[CrossRef\]](#)
29. Zdenek, K.; Hainan, C.; Xiaowei, L. Transfer learning and deep convolutional neural networks for safety guardrail detection in 2D images. *Autom. Constr.* **2018**, *89*, 58–70. [\[CrossRef\]](#)
30. Seo, J.; Han, S.; Lee, S.; Kim, H. Computer vision techniques for construction safety and health monitoring. *Adv. Eng. Inform.* **2015**, *29*, 239–251. [\[CrossRef\]](#)
31. Weili, F.; Lieyun, D.; Hanbin, L.; Peter, E.L. Falls from heights: A computer vision-based approach for safety harness detection. *Autom. Constr.* **2018**, *91*, 53–61. [\[CrossRef\]](#)
32. Nevada Department of Transportation. *Risk Management and Risk-Based Cost Estimation Guidelines*; Nevada Department of Transport: Carson City, NV, USA, 2012.
33. Kim, J.M.; Woods, P.K.; Park, Y.J.; Kim, T.; Son, K. Predicting hurricane wind loss by claim payout based on Hurricane Ike in Texas. *Geomat. Nat. Hazards Risk* **2015**, *7*, 1513–1525. [\[CrossRef\]](#)
34. Ahmed, S. Causes and effects of accident at construction site: A study for the construction industry in Bangladesh. *Int. J. Sustain. Constr. Eng. Technol.* **2019**, *10*, 18–40. [\[CrossRef\]](#)
35. Moyce, S.C.; Schenker, M. Migrant workers and their occupational health and safety. *Annu. Rev. Public Health* **2018**, *39*, 351–365. [\[CrossRef\]](#)
36. Sousa, E.; Agudelo-Suárez, A.; Benavides, F.G. Immigration, work and health in Spain: The influence of legal status and employment contract on reported health indicators. *Int. J. Public Health* **2010**, *55*, 443–451. [\[CrossRef\]](#)
37. Lim, B. Comparison of Occupational Accidents According to the Type of Employment of Safety and Health Managers. Linked to Public Works: Korea Occupational Safety and Health Agency Published Data. 2015. Available online: <http://oshri.kosha.or.kr/oshri/publication/researchReportSearch.do?mode=view&articleNo=63449&article.offset=0&articleLimit=10&srSearchVal=%EA%B3%A0%EC%9A%A9%ED%98%95%ED%83%9C> (accessed on 2 February 2021).
38. Gambatese, J.A.; Behm, M.; Hinze, J.W. Viability of designing for construction worker safety. *J. Constr. Eng. Manag.* **2005**, *131*, 1029–1036. [\[CrossRef\]](#)
39. Cho, J. A study on the cause analysis and preventive measures for each type of disaster at the construction site. *J. Korea Saf. Manag. Sci.* **2012**, *14*, 7–13.
40. Kwon, S.C. Use of frequency analysis of exposure of hazards by occupations: Findings from the third and fourth Korean working conditions survey. *Soonchunhyang Med. Sci.* **2019**, *25*, 37–45. [\[CrossRef\]](#)
41. Na, H.; Park, B.H. Developing accident models of rotary by accident occurrence location. *Int. J. Highw. Eng.* **2012**, *14*, 83–91. [\[CrossRef\]](#)
42. Ajayi, A.; Oyedele, L.; Owolabi, H. Deep learning models for health and safety risk prediction in power infrastructure projects. *Risk Anal.* **2020**, *40*, 2019–2039. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Gu, J.; Wang, Z.; Kuen, J. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, *77*, 354–377. [\[CrossRef\]](#)
44. Zhong, G.; Wang, L.N.; Ling, X.; Dong, J. An overview on data representation learning: From traditional feature learning to recent deep learning. *J. Financ. Data Sci.* **2016**, *2*, 265–278. [\[CrossRef\]](#)
45. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Carreira-Perpinan, M.A.; Hinton, G.E. On contrastive divergence learning. *Aistats* **2005**, *10*, 33–40.
47. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. Available online: <http://dl.acm.org/citation.cfm?id=2999134.2999257> (accessed on 18 September 2021). [\[CrossRef\]](#)
48. Bae, S.W.; Yoo, J.S. Apartment price estimation using machine learning: Gangnam-gu, Seoul as an example. *Real Estate Stud.* **2018**, *24*, 69–85.
49. Li, Z.; Gong, B.; Yang, T. Improved dropout for shallow and deep learning. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, 2523–2531.
50. Ryu, J.D.; Park, S.M.; Park, S.H.; Kwon, C.W.; Yoon, I.S. A study on the development of a model for predicting the number of highway traffic accidents using deep learning. *J. Korean ITS Soc.* **2018**, *17*, 14–25.
51. Kim, J.M.; Bae, J.; Son, S.; Son, K.; Yum, S.G. Development of model to predict natural disaster-induced financial losses for construction projects using deep learning techniques. *Sustainability* **2021**, *13*, 5304. [\[CrossRef\]](#)
52. Bae, J.; Yum, S.G.; Kim, J.M. Harnessing machine learning for classifying economic damage trends in transportation infrastructure projects. *Sustainability* **2021**, *13*, 6376. [\[CrossRef\]](#)
53. Kingma, D.P.; Ba, J. ADAM: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference for Learning Representations, San Diego, CA, USA, 30 January 2015; Volume 9, pp. 1–15.