ISSN:1738-6764, eISSN:2093-7504

Development of AI-Based Water Pipe Aging Condition Evaluation Algorithm

Gi-Tea Kim [1], Byung-Won Min [2] and Yong-Sun Oh [3,*]

[1] Division of Information and Communication Convergence Engineering, Mokwon University; Ph.D. completion; gitae@kwater.or.kr
[2] Division of Information and Communication Convergence Engineering, Mokwon University; Professor; minfam@mokwon.ac.kr
[3] Division of Information and Communication Convergence Engineering, Mokwon University; Professor; ysunoh@mokwon.ac.kr
* Correspondence

https://doi.org/10.5392/IJoC.2022.18.4.109

Manuscript Received 01 September 2022; Received 20 December 2022; Accepted 20 December 2022

Abstract: *As interest in the development of artificial intelligence technology in the water supply field increases, an artificial neural network algorithm that can predict improved decision ratings through repetitive learning using results of aging pipe condition evaluation data should be developed and the most reliable prediction model should be presented through a verification process. An algorithm was developed to predict pipeline ratings by updating weights through backpropagation so that 12 items of indirect evaluation data according to the 2020 Han River Basin's basic plan could be pre-processed, such as standardizing input values and applying artificial neural network algorithms. As a result of algorithm accuracy verification, if there is sufficient data and the repetitive learning and upgrades are continuously conducted, the prediction accuracy will become higher and a reliable AI-based water supply pipe condition evaluation model to be used nationwide will be developed in the future.*

Keywords: Artificial Neural Network; Data Preprocessing; Pipeline Condition Evaluation; Algorithm Model; Learning Process; Indirect Evaluation; Direct Evaluation

1. Introduction

1.1 Background and Purpose of Study

For water pipes supplying tap water produced at a water treatment plant, accurate decisions are needed to replace, repair, and clean the facilities in a timely manner for their improvement to prevent accidents related to water quality and leakage, and for that purpose, it is significantly important to produce an accurate estimation of their status.

The most commonly-used method for such improvement decisions by analyzing the status of dilapidation of water supply pipelines is an estimation of the residual strength of the residual thickness for a metal pipe based on the prediction of the depth of corrosion, and a calculation of safety factor(SF) reflecting the load coming from internal and external pressures depending on the burying conditions. However, it is causing trials and errors in the field as most of the studies are being carried out with limited data, such as investigation data of a specific section and accident history data, and are difficult to implement direct verifications.

Under the circumstances, it is necessary to develop various prediction models per purpose of application and to study reliable and accurate decision-making methods.

1.2 Content and Scope of Study

This study is aimed to develop an algorithm model using an artificial neural network that can predict the grades of pipelines required to be determined for improvement through repeated learning based on the results of dilapidation status evaluation data of old water supply pipes, and to present the most reliable prediction model by carrying out the verification process.

To this end, this study was conducted with the indirect and direct status evaluation data of the old pipelines managed by seven local governments along the Han River, which had been collected in 2020.

The prediction procedures of the algorithm model in this study are illustrated in Figure 1. It collects indirect and direct investigation data on the assessed levels of dilapidation, pre-processes the collected data, conducts repeated learning, and establishes an algorithm model of artificial neural network.

In addition, it is intended to offer the most reliable decision-making algorithm model for improvements of pipelines by going through the process of predicting and verifying the built model.

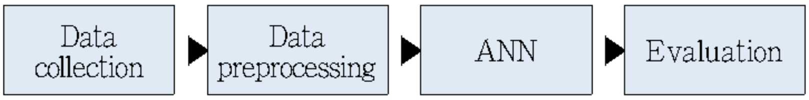


Figure 1. Prediction model development procedure

2. Decision-making Method Based on Existing Status Evaluation

In Korea, most of the status evaluations of old water pipelines are done utilizing a safety factor, a methodology for decision-making for improvements from a physical point of view, and Figure 2 shows a brief description of its procedures[1].

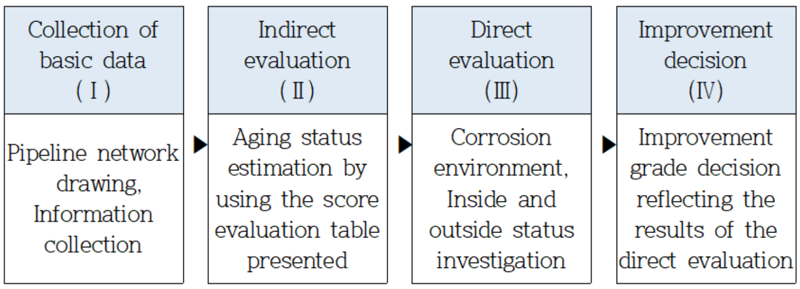


Figure 2. Dilapidation status evaluation procedure

The first stage involves a collection of basic data of pipelines and the division of pipelines by block and pipe type. After that, pipelines are categorized according to their characteristics by grouping the similar subdivided pipelines.

The indirect evaluation done in the second stage entails a prediction of the current aging status by using the score evaluation table presented for each pipe type and produces results in the form of three grades from I to III, and a section rated as Grade II and involved in water quality complaints and a section rated as Grade III are selected for direct evaluation.

The direct evaluation in the third stage is to select an on-site investigation spot based on the indirect evaluation result (grade) and directly investigate the operation/burial conditions and the degrees of deterioration of the inside and outside of the pipe so as to conduct an intensive investigation into 21 items by collecting a sample of the pipe and carrying out an analysis. In the last fourth stage, the grade and method for improvement are decided reflecting the results of the direct investigation and the criteria for improvement decision.

3. Development of Algorithm for Status Evaluation using Artificial Neural Network

3.1 Overview of Algorithm Model of Artificial Neural Network

An artificial neural network is an analysis model that refers to a process of finding weights of complex functional formulas that clearly explains the relationship between input and output[2].

For this, the linear sum of the input variables of the pipeline data is put into the activation function (ReLU, Softmax, etc.) and optimized through repeated learning to make the rate of concordance between the calculated result and the actual value (direct evaluation result) higher than 90%. At this time, the weight is updated through the backpropagation process to reduce the error and ultimately to predict the pipeline grade[3, 4].

An artificial neural network consists of an input layer, an intermediate layer, an output layer and each weight that connects an input layer with an intermediate layer, and an intermediate layer with an output layer[5, 6].

As the structure of the artificial neural network is set in 12 items of input layers and 5 grades of output layers, the number of hidden layers and hidden nodes was estimated after numerous tests as one hidden layer with the best prediction performance with 100 hidden nodes as shown in Figure 3.

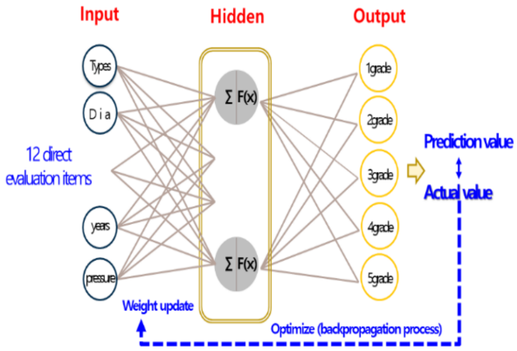


Figure 3. Prediction model ANN structure

Table 1 compares the existing indirect evaluation method for pipelines condition evaluation and the developed artificial neural network model method.

Table 1. Comparison of traditional and AI methods

	Existing Indirect Evaluation	AI Evaluation
Weighted value	Score evaluation table	Initial weight(He)
Score calculation	Score calculation method	Update weighting factors through backpropagation
Learning	Update Score evaluation table	Backpropagation algorithm (Update initial weight)
Verification	Status evaluation grade table	Predict ratings by reducing error values

3.2 Data Preprocessing and Grade Prediction Method

To develop an artificial-neural-network-based algorithm model, data pre-processing was first conducted, and data transformation was carried out by extracting, recombining, and standardizing the input value formats of 12 indirect evaluation items(as shown in Table 2) required for pipeline grade prediction.

Table 2. 12 indirect evaluation items

	Items
Pipeline data	Types, Diameter, Years
Corrosion environment	Soil Types, Water corrosiveness, Soil resistivity, CP system
Inside and outside load	Maximum pressure, Depth of burying, Road Types
Accident history	Number of breakages
Joint quantity	Valve/Water pipe quantity

As the input values of the artificial neural network (ANN) used for pipeline grade prediction can only be numeric data, all character data must be converted to numeric. Furthermore, blanks in character data are removed and data values with the same meaning for each region are unified into one.

As for pipe types, pipe types other than CIP, DCIP, and SP are classified as non-metallic. Among the 12 pipeline data items, as the data regarding the pipe types, types of soil, and road types are in characters, data preprocessing, including standardization of input value format, was done to apply the artificial neural network algorithm as shown in Table 3. In addition, data outliers and omissions were removed[7].

Table 3. Pipeline data standardization

Pipe types	CIP	DCIP	SP	Non-metallic
Conversion	0	1	2	3
Road types	Industrial road	Walking road	One-lane road	Two-lane road
Conversion	0	1	2	3
Soil Types	Sand	Sand and soil	Gravel	Gravel and soil
Conversion	0	1	2	3

Table 4. lists up the actual input values of 12 pipeline data entered to the artificial neural network algorithm. Results of the grade evaluation for a decision to improve a pipeline are classified into five grades.

Table 4. Data actual inputs

Pipe types	Diameter	Years	Soil Types	Water corrosiveness	Soil resistivity	Depth of burying	Road types
DCIP	100	24	Sand	-1.1	11638	1.4	2 lanes
DCIP	200	21	Sand	-1.1	11638	1.5	2 lanes
DCIP	100	21	Sand	-1.1	11638	1.1	1 lane
DCIP	200	21	Sand	-1.1	11638	1.2	1 lane
DCIP	100	24	Sand	-1.1	11638	1.4	2 lanes

4. Application of Artificial Neural Network Algorithm Model

4.1. Learning and Verification of Algorithm Model

Data targeted in this study included indirect and direct evaluation data(as shown in Table 5) of pipelines for water conveyance, water supply, and drainage in three cities in Province A and four cities and counties in Province B(as shown in Table 6) involved in maintenance of aged water pipelines carried out in 2020[8].

Table 5. Indirect and direct status evaluation data

	Sum	Province A	Province B
Data lines	1626	736	890

Table 6. Pipeline data

	DCIP	SP	Non-metallic	Etc.
Province A	357	282	69	28
Province B	11	106	759	14

The data was acquired in the form of an Excel file, with the evaluation data of one pipeline per line. For this, data standardization was implemented and the entire data of Province A and Province B are divided into 60% (learning): 20% (validation): 20% (evaluation) for learning, verification, and evaluation(as shown in Table 6) of the artificial neural network algorithm model ultimately to conduct repeated learning and to predict a grade for a decision for improvement.

Table 7. Split data ratio

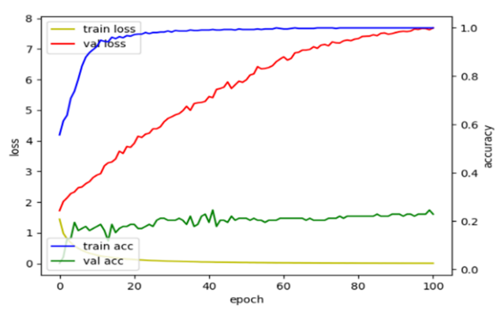
	Ratio	Purpose
Learning	60%	Model Learning
Validation	20%	Model Check
Evaluation	20%	Model Verification

To form a normal model by evenly distributing the data when learning artificial neural network data, the data was first randomly mixed and divided into 60:20:20 using the Scikit-learn library for data classification.

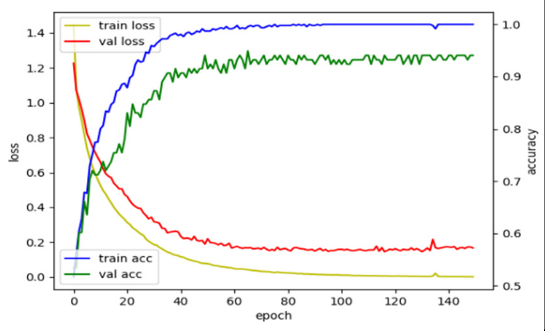
The input, output, and hidden layers of the artificial neural network were generated utilizing the Tensorflow-based Keras library.

For learning, 60% of the data was applied for iterative learning, and as for the model validation of the data, iterative learning was implemented to make sure that the rate of concordance between the prediction result produced by the algorithm model and direct evaluation value was higher than 90 percent by checking the graph in which Val Loss gradually decreased and Val Acc gradually increased, as shown in Fig 4[9].

The number of hidden layers and the number of hidden nodes were sequentially reduced, and tests were conducted to confirm the shape and accuracy of the model. In the test of the model, the shape of the model was confirmed by proceeding with 3 hidden layers and 100 hidden nodes, respectively. As the Val Loss value increased abnormally, the model was broken. The same was true when there were two hidden layers.



Abnormal training graph



Normal training graph

Figure 4. Training graph

4.2 Evaluation of Algorithm Model

For better accuracy of the predictive evaluation of the artificial neural network algorithm model, the outliers that aggravated the performance of the model were removed after looking into the input values, such as the age and burial depth of pipes, among the status evaluation data of Province A and B.

IQR (Inter Quantile Range) method was chosen as the method of removing outliers[10].

A predictive evaluation was carried out with 20 percent (for evaluation) of the data coming from three cities in Province A and four cities and counties in Province B using the artificial neural network algorithm model learnt from their pipeline status evaluation data to check the reliability of the developed model.

The verification results of the prediction capability of the developed algorithm model indicate that Province A showed a high accuracy of 94.9%, while Province B displayed an accuracy of 81.9% as shown in Table 8.

Table 8. Accuracy prediction result

	Province A	Province B
Accuracy	129ea/136ea (94.9%)	145ea/177ea (81.9%)

Province B was confirmed to have a slightly lower accuracy than Province A, but this is presumably attributed to lack of learning as the data of Province B covered a wider area and involved more non-metallic pipes disadvantageous in the existing aging status evaluation, which can be rectified if more data are included in the learning process.

5. Conclusions

In this study, the direct and indirect evaluation data based on the basic plan for maintenance of old pipelines were applied to the artificial neural network algorithm model

After developing an artificial neural network algorithm model using the dilapidation status evaluation data of three cities in Province A and four cities and guns in Province B in 2020, the predictions of the grades of old pipelines for a decision for improvement indicated that Province A showed an accuracy of 94.9 %, while Province B recorded an accuracy of 81.9 %.

Province B was confirmed to have a slightly lower accuracy than Province A, but this is presumably attributed to lack of learning as the data of Province B covered a wider area and involved more non-metallic pipes disadvantageous in the existing aging status evaluation, which can be rectified if more data are included in the learning process.

Against this backdrop, it was confirmed that the accuracy of predictive evaluation increased only when the status evaluation data of the old pipelines was continuously secured and the data showing that the evaluation grades of all pipe types were evenly distributed was repeatedly learned.

In the future, in case a variety of, and a larger quantity of data is secured from all over the country and repeatedly learnt, the reliability of pipe grade prediction using the artificial neural network will be improved, which in turn is expected to play a crucial role in the decision making process for improvements when applied in the predictions of the dilapidation status of the pipelines nationwide.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Water Pipeline Precision Investigation Manual(Draft), Ministry of Environment, pp. 10-24, 2020.

[2] Y. R. Choi and K. H. Kim, “Artificial Intelligence Overview and Application Cases,” ie Magazine, p. 23, 2016.

[3] Y. I. Cho, M. J. Oh, Y. S. Seok, S. J. Lee, and M. I. Roh, “Resistance Estimation of a Ship in the Initial Hull Design Using Deep Learning,” Korean Journal of Computational Design and Engineering, pp. 203-210, 2019. doi: 10.7315/CDE.2019.203

[4] Cybenko, George. "Approximation by Superpositions of a Sigmoidal Function," Mathematics of Control, Signals, and Systems, pp. 303-314, 1989. doi: 10.1007/BF02551274

[5] S. R. Jo, H. N. Sung, and B. H. Ahn, “A Comparative Study on the Performance of SVM and an Artificial Neural Network in Intrusion Detection,” Journal of the Korea Academia-Industrial cooperation Society, pp. 703-711, 2016. doi: 10.5762/KAIS.2016.17.2.703

[6] C. W. Lee, S. J. Hong, J. S. Park, and D. G. Yoo, “Assessment of Water Supply Pipes Deterioration Based on Deep Learning,” Journal of the Korean Society of Civil Engineers, pp. 160-161, 2020

[7] K. H. Han, Y. J. Ryu, T. S. Kim, and J. H. Heo, “Input Variables Selection of Artificial Neural Network Using Mutual Information,” Korea Water Resources Assocition, pp. 85-87, 2010. doi: 10.3741/JKWRA.2010.43.1.81

[8] Korea Water Resources Corporation, "A Report on the schematic plan for the Improvement of the Aged Water Pipeline," 2021.

[9] J. Y. Seo, “Deep Learning TensorFlow Textbook,” Gilbut, p. 25, 2021.

[10] H. K. Park, J. S. Lee, W. H. Song, and S. S. Park, “Technology Forecasting using Patent Outlier Remova,” Journal of Korean Institute of Intelligent Systems, pp. 177-184, 2018. doi: 10.5391/JKIIS.2018.28.2.177