치의학박사 학위논문

Novel method to diagnose extraction patterns with the artificial intelligence decision-making model using neural network

신경망 인공지능 의사결정 모델을 이용한 발치 진단의 새로운 방법 제안

2016년 2월

서울대학교 대학원

치의과학과 치과교정학 전공

정 석 기

치의학박사 학위논문

Novel method to diagnose extraction patterns with the artificial intelligence decision-making model using neural network

지도교수 김 태 우

이 논문을 치의학박사 학위논문으로 제출함

2015년 10월

서울대학교 대학원

치의과학과 치과교정학 전공

정 석 기

정석기의 치의학박사 학위논문을 인준함

2015년 12월

위 원 장 ____________________ (인)

부 위 원 장 ____________________ (인)

위 원 ____________________ (인)

위 원 ____________________ (인)

위 원 ____________________ (인)

Abstract

Novel method to diagnose extraction patterns with the artificial intelligence decision-making model using neural network

Seok-Ki Jung, DDS, MSD

Department of Orthodontics, Graduate School, Seoul National University
(Directed by Professor **Tae-Woo Kim**, *DDS, MSD, PhD)*

Introduction: The diagnosis of extractions in the orthodontic treatment is important and difficult, because that decision has tendency to be based on the practitioners' experiences. The purpose of this study was to construct an artificial intelligent expert system for the diagnosis of extraction using neural network machine learning (NNML) and to evaluate performance of this model.
Methods: The subjects consisted of 156 patients in total. Input data consisted of 12 cephalometric variables and additional six indices. Output data consisted of three bits to divide extraction patterns. Four NNML models for the diagnosis of extractions were constructed using backpropagation algorithm, and were evaluated.
Results: The success rates of the models showed 93% for the diagnosis of extraction versus non-extraction, and showed 84% for the detailed diagnosis of the extraction patterns.
Conclusions: This study suggests that artificial intelligent expert systems using neural network machine learning could be useful in orthodontics. Improving performance was achieved by the components such as proper selection of the input data, appropriate organization of the modeling, and preferable generalization.

Key words: Machine learning, Extraction, Diagnosis, Neural network
Student number: 2014-30728

Novel method to diagnose extraction patterns with the artificial intelligence decision-making model using neural network

Seok-Ki Jung, DDS, MSD

Department of Orthodontics, Graduate School, Seoul National University
(Directed by Professor **Tae-Woo Kim***, DDS, MSD, PhD)*

신경망 인공지능 의사결정 모델을 이용한 발치 진단의 새로운 방법 제안

서울대학교 대학원 치의과학과 치과교정학 전공
(지도교수 : 김 태 우)

정 석 기

- Contents -

I. Introduction

The most important thing in the orthodontic treatment is to decide treatment plan.[1] Furthermore, the most important thing in the treatment planning is the decision-making of extractions and the teeth that would be extracted. It is because extractions are the irreversible procedures. Therefore a prudent decision of extractions is required. If a wrong decision has been made, many problems could arise in all time of the orthodontic treatment. Undesirable results could be come out or the treatment could not be finished in the worst case. These problems appear as the failure of the anchorage control, abnormal inclination of the anterior teeth, unfavorable profile, improper occlusion, inadequate overjet and overbite, and difficulties of the closure of extraction spaces. Generally, most orthodontists make a decision with data from the clinical evaluation, photographs, dental models, and radiographs based on their experiences and knowledge. Since there is no formula for the treatment planning, a decision depends on the practitioners' heuristics in many cases.[2] This often cause intra- and inter-clinician reproducibility of the treatment planning process.[3] In addition, different records which was used for diagnosis can make differences of the treatment plan.[4-6] Moreover, differences of the treatment planning could occur between experienced practitioner and less-experienced practitioner.[7] In particular, differences of extractions could be critical. If it is possible to share decisions of the experienced practitioner, that would be very helpful. However, decisions cannot be standardized with such as combination of measurements. Thus another approach is needed.

Recently, there are many studies about artificial intelligence and bioinformatics.[8-10] One of those is the machine learning using neural network system.[11,12] This emulates human learning in the situation which cannot be formulated or standardized. Human neural system consists of neurons. These neurons are linked at the synapse in order to send information. By repeated learning, each synapse linkage can be reinforced or weakened. Neural network system is formed through these procedures, and answers for afterward questions can be derived from this system. In the machine learning using neural network, there are neurons linking from the input to the output, and each neuron is linked at the synapse. In each synapse, information of the input neurons is collected by weighting technique. Weighted values are adjusted through iterative learning (Fig 1). Excessive iterative learning can elevate goodness of fit of the training set. However, errors of the other set can be also increased. It is called the overfitting. In order to avoid this, validation set is introduced in order to stop learning and to make a generalized model (Fig 2). The generalized decision-making model can be formed through these procedures.

The aim of this study was to make an artificial intelligent decision-making model for the diagnosis of extractions using neural network machine learning. In addition, it was supposed to evaluate validity and accuracy of this model.

II. Review of Literature

Introduction about machine learning

For the decision-making problems, a number of expert systems have been developed in the field of engineering science. [8,9,12-14] Machine learning is the method to find answers using artificial intelligence. A classification problem is one of decision-making problem that can be applied with machine learning.

If there is the set of questions and answers, machine learning system learns the set iteratively. Through this procedure, a system adjusts its own program to make goodness-of-fit better using various techniques and algorithms. There are many algorithms such as decision-tree learning, association rule learning, genetic algorithms, inductive logic programming, support vector machines, clustering, bayesian networks, reinforcement learning, representation learning, similarity and metric learning, sparse dictionary learning, and artificial neural networks. [8,9,11,12,15]

Algorithms can be different but the purpose is to establish the system through iterative learning of the set, called training set. The final aim is making a good system to solve another set as well, called test set.

The difference between the percentage of correct answers for the training set and the test set should be minimal. As the difference bigger, that system is not a well generalized system. It means the system specialized in the training set only. It can give the right answers for the training set but cannot give the right answers for the other set. It is also called overfitting.

To avoid overfitting, the validation set is needed.[16] The validation set can give information of when the iterative learning should be stopped. At the minimal point of the percentage of correct answers for the validation set, the learning is stopped and the system is established.

With the artificial intelligence system being obtained through machine learning, we can get the answers for many problems, especially the problem which cannot be solved easily. The vague problem is one of reasons using artificial intelligence system such as machine learning.

Artificial neural network systems

Artificial neural network is one of algorithms to construct machine learning system.[8,9,11,17] It imitates the human neural network system. The human neural network system consists of neurons. Each neuron is linked in the synapse and transfers information each other. This information is integrated through the complex neural network. Some information is strengthened and other information is weakened.

Artificial neural network is comprised of numerous nodes and their linkages. [16] A node plays a role in each neuron and a linkage plays a role in each synapse. There are three layers in the artificial neural network system. The first is the input layer which receives input signal. The second is the output layer which makes a result signal. The last is the hidden layer which mediates information from input layer to output layer. Nodes in the hidden layer mimic intermediate neurons. An artificial neural network is an interconnected group of nodes.

Machine learning is performed based on this artificial neural network. When transferring information from one node to another node, it is multiplied with the weighted value. All of information transferred to one node is integrated and compared with the true value.[16] According to the result, adjustment of the weighted value is performed using backpropagation algorithm. Throughout iterative learning, all of the weighted values are adjusted and goodness-of-fit of the training set can be elevated.

Compared to other algorithms for machine learning, artificial neural network has advantages of easiness for application. Meanwhile, artificial neural network is used mainly for the classification problems. In this study, decision-making for the diagnosis of extractions was a kind of classification problem. Thus, the artificial neural network algorithm was selected for machine learning.

Expert systems in orthodontics

There have been a number of published articles on the development of the decision-making systems, such as the decision tree based,[18,19] fuzzy logic based,[13,20] template matching based,[21] and neural network based system.[17]

Hicks et al[22] stated that our brains use two modes of reasoning: heuristic (intuitive, automatic, implicit processing) and analytic (deliberate, rule-based, explicite processing). Since choice heuristics were often biased by prior evaluations and preferences for the alternatives being considered. Clinical orthodontics could be relevant with common cognitive biases.

Poon et al[23] suggested a new approach to knowledge acquisition known as Ripple-Down-Rules for the development of expert system in clinical orthodontics. This system had a knowledge base of six hundred and eighty rules.

Brickley et al[11] introduced the concepts of neural network based system in dentistry. They stated that neural networks might become important decision making tools and could have applications within dentistry.

Lux et al[24] analysed the growth of 43 orthodontically untreated children. Through the use of an artificial neural network, namely self-organizing neural maps, the resultant growth data were classified and the relationships of the various growth patterns were monitored.

Sims et al[14] reported that FRIL (Fuzzy Relational Inference Language) could be used as the powerful tool to make an expert system for non-specialist dentists to solve orthodontic problems. The use of fuzzy relations and descriptors could be useful for modeling of orthodontic diagnosis processes.

Stephens et al[19] reported about the validation of an rule-based orthodontic expert system for fixed appliance treatment planning. The program used expert system techniques including rule-based reasoning and fuzzy logic-based representations of orthodontic knowledge. The treatment plans generated by an expert system were judged to be of similar quality to those of the orthodontist.

Yagi et al[25] suggested a decision-making system for orthodontic treatment planning based on direct implementation of expertise knowledge, and demonstrated the prediction accuracy of 90.5% for decision-making process regarding tooth extraction.

Noroozi[20] suggested an orthodontic treatment planning software. This software could receive patient data in both graphic and numeric forms and propose a treatment plan for nonsurgical orthodontic patients.

Hammond et al[21] stated that traditional rule-based expert system had some limitations when applied to orthodontic diagnosis and treatment planning and these limitations might be avoided by using a case-based system.

Takada et al[26] reported the mathematical model to simulate whether or not to extract teeth in optimizing orthodontic treatment outcome. The optimum decision of whether or not to extract teeth was predicted by means of a template-matching technique with nearest neighbor search. The success rate of 90.4% was shown at its prediction performance.

Xie et al[17] reported an artificial neural network (ANN) modeling for deciding extractions for orthodontic treatment. A 23-13-1 Back Propagation (BP) ANN model was constructed to determine whether extraction was needed. The result was 100% for the training data set, and the 80% correct for the test data set.

Yu et al[27] used a machine learning technique for evaluation of facial attractiveness for patients with malocclusion. A support vector regression (SVR) function was set up according to the coordinate values of landmarks. Although some ratios and angles were found to have a close correlation with facial attractiveness, they could not be used for comprehensively evaluation for facial attractiveness from a set of orthodontic photographs.

Moghimi et al[28] used a hybrid genetic algorithm and artificial neural network (GA-ANN) system for predicting the sized of unerupted canines and premolars during the mixed dentition period. The prediction error rates using the hybrid GA-ANN algorithm were smaller than those using linear regression analyses.

Nieri et al[15] applied the Bayesian network to evaluate the relative role and possible causal relationships among various factors affecting the diagnosis and final treatment outcome of impacted maxillary canines. Bayesian network analysis was useful to identify possible relationships among the variables considered for diagnosis and treatment of impacted canines.

Previous many articles have already studied the necessity and the possibility of the decision-making expert system in orthodontics. However, there were no studies about expert system for the diagnosis of extractions according to the extraction pattern.

III. Material and Methods

The subjects consisted of 156 patients who had visited Seoul National University Dental Hospital, Seoul, South Korea for orthodontic consultation. Exclusion criteria were persons who had had unerupted permanent teeth or missing teeth (except for the third molars), malformed teeth, previous orthodontic treatment history, maxillofacial deformities, and orthognathic surgery. Inclusion criteria were persons who was included in five treatment plan groups as follows: non-extraction, maxillary and mandibular first premolar extractions (Ext44), maxillary and mandibular second premolar extractions (Ext55), maxillary first premolar and mandibular second premolar extractions (Ext45), and maxillary first premolar extractions only (Ext40) (Table I). For all samples, treatment plans were decided by one orthodontic specialist who had experienced more than 10 years.

Lateral cephalograms were filmed as orthodontic records for all samples. All tracings were performed by single investigator and repeated twice with interval of two weeks in order to analyze measurement errors. The reference points were digitized by V-ceph program (ver 5.3, Osstem Inc., Seoul, Korea). Twenty-six landmarks and 12 measurements were chosen (Fig 3).

With this sample, 96 persons were assigned to the learning set and 60 persons rest were assigned to the test set (Table I). The test set was used only for evaluation of the models. Two-thirds of the learning set was assigned to the training set and one-third of the learning set was assigned to the validation set. In order to find the optimal model, sliding window validation was performed. Sliding window validation is the validation technique to choose validation set through the window moving sideways from the serial data.[16] In order to avoid the overfitting, iterative learning was stopped at the minimum error point of the validation set. Next, through the evaluation for the test set, the adequacy and the accuracy were evaluated and the best-fit model was chosen.

Two-layer neural network including one hidden layer was selected for the machine learning. There were four hidden nodes in the hidden layer. Hidden nodes play a role of interneuron in the artificial neural network system, and learning is performed through their weighted values adjustment. Twelve measurements was selected for the input data as follows: ANB angle, Overjet, Björk sum, Overbite, Upper 1 to SN angle, Upper 1 to Occlusal Plane angle, IMPA, Lower 1 to Occlusal Plane angle, Interincisal angle, Upper lip to E-line, Lower lip to E-line, and Nasolabial angle. They had clinical relevance of such as anteroposterior relationship, vertical relationship, teeth inclination and soft tissue characteristics, respectively. In addition, six indices - maxillary arch length discrepancy index, mandibular arch length discrepancy index, molar key index, large overjet index, protrusion index, and chief complaint index for protrusion - were included into the input data (Table II). Input data consisted of total eighteen elements with this manner. Max-min normalization was chosen for the normalization of the input data in the range of 0 to 1. Learning rate was 0.9 and sigmoid function was chosen as the activation function. Language R-program (http://www.r-project.org/) was used for coding in order to construct machine learning models.[29] Backpropagation algorithm was applied to adjust weighted values.

Output data were composed of the three bits. Dx_ext was the index about whether to need extractions. The value of 0 meant non-extraction and the value of 1 meant extraction. Dx_diff was the index about whether to need the differential extraction between maxillary and mandibular arch. The value of 0 meant identical extraction such as Ext44 and Ext55. The value of 1 meant differential extraction such as Ext45 and Ext40. Dx_more was the index about whether to need more retraction. The value of 0 meant mild-to-moderate retraction case such as Ext55 and Ext45. The value of 1 meant moderate-to-severe retraction case such as Ext44 and Ext40 (Table III).

Trainings were performed with three stages, and four best-fit models were selected through those trainings. The first classifier (Classifier_1) was the model deciding to extract or not, which output was Dx_ext. The second classifier (Classifier_2) was the model deciding differential extractions or not, which output was Dx_diff. The third stage was for making the models deciding more retraction or not, which output was Dx_more. In the third stage, two classifiers (Classifier_3 and Classifier_4) regarding identical and differential extractions were derived (Fig 4). Extraction diagnosis of total data was performed by constructed classifiers. In comparison with actual diagnosis, decision-making success rates of Dx_ext, Dx_diff, and Dx_more were calculated. Finally, total success rate of the diagnosis of extractions was calculated.

IV. Results

The results of decision-making success rates were summarized in Table IV. In addition, each learning and validation curve was shown in Fig 5.

The intraclass correlation coefficient (ICC) was used to evaluate the test-retest reliability of the tracings and its values were scored as follows: ICC < 0.4, poor reliability; $0.4 <$ ICC < 0.75, moderate reliability; ICC > 0.75, excellent reliability.[30] The ICC values in this study ranged from 0.97 to 0.99, demonstrating the excellent reliabilities.

In the diagnosis of extraction versus non-extraction, decision-making success rates were 92% in the training set, 94% in the validation set, 93% in the test set, and 93% in total. In the diagnosis of identical versus differential extraction, success rates were 88% in the training set, 100% in the validation set, 85% in the test set, and 89% in total. In the diagnosis of more retraction in identical extraction, success rates were 88% in the training set, 75% in the validation set, 85% in the test set, and 84% in total. In the diagnosis of more retraction in differential extraction, success rates were 95% in the training set, 100% in the validation set, 95% in the test set, and 96% in total. Through the sequential application of decision-making models, final success rates were 85% in the learning set, 82% in the test set, and 84% in total (Fig 6).

In the analysis of the failed diagnosis cases, seven cases were reversed between Ext44 and Ext45, which was the biggest portion. Next, six cases were reversed between Ext 55 and non-extraction. In the total 25 cases of the failed diagnosis, unacceptable decisions were in four cases only. Decisions for other cases were acceptable because they had been borderline cases. Including these cases, decision-making success rates rose as 97%.

V. Discussion

For the classification problems, machine learning has been used in many studies.[24-26,31-33] The diagnosis of extractions can be approached as a kind of classification problem. Takada et al had reported decision-making system for orthodontic treatment planning using template-matching based system.[26] Template-matching means finding similar case from the established database, which is the different method with this study. Similar to this study, Xie et al had used artificial neural network modeling for determining extraction or non-extraction.[17] The previous study had determined necessity for extraction only. However, this study determined extraction positions in addition. Furthermore, decision-making success rates were improved certainly. In the previous study, success rates were 100% in the training set, and 80% in the test set. The difference of the success rates between the training set and the test set could mean overfitting. In order to minimize overfitting and to verify fitness of the model, samples were divided into the learning set and the test set from the beginning in this study. In addition, the learning set was divided into the training set and the validation set in order to make a generalized model. As a result of this, success rates of the training set, the validation set, and the test set were similar in this study. It implies that the model of this study was generalized better.

In order to treat skeletal Class III patients, surgical orthodontic treatments are preferred rather than camouflage treatments to make an ideal result. Thus, the diagnosis of extractions was limited into five patterns in this study, because it could cover the most cases of the orthodontic treatment only.

The main reasons for extractions are crowding and protrusion.[34] In order to reflect this, the indices of arch length discrepancy and protrusive profile were added. Through the pilot study, grouped index showed better performance than numerical value itself. The reason might be that group itself had been more important for the decision of extractions. Chief complaint index for protrusion was added, because it could affect the diagnosis in the borderline cases. Lastly, molar key index and large overjet index were added, because they were important components for the diagnosis of differential extractions. For these reasons, six additional indices were added into the input data.

Output data were three bits, and learning was performed through four steps in this study. It is because using output of 0 or 1 showed better performance in the pilot study. Therefore, two bits of output were needed for four cases of extraction diagnosis patterns. Another bit was needed for division of extraction and non-extraction.

Though step-by-step learning had shortcomings of accumulating the errors, the goodness of fit was better than one-step learning. It was because simpler system might have higher success rate. In addition, the case that failed previous step tended to fail also with next step. Thus accumulation of the errors could be minimized.

The limitation of this study was that the diagnosis of extractions was confined to the non-surgical procedure. In addition, the model could not cover the cases such as missing teeth, uncommon extraction, asymmetry, and soft tissue functions. Further study for the diagnosis of surgical procedures and other cases will be planned. Through this, a complete model that covers all cases could be established. Other limitation of this study was the ambiguity of the protrusion index. The protrusion index could be a little subjective. It was difficult to express exactly protrusion by the combination of several measurements. However, if the protrusion index is applied consistently, the model could make a reasonable result. If necessary, customized diagnostic learning for each practitioner could be also possible. That will reflect the practitioner’s preference.

In fact, there is no correct answer for the diagnosis of extractions. The aim of this study was not for finding a right answer. With mimicking decision of experienced experts, artificial intelligent expert system could give a reference to the less experienced practitioners. Clinicians can choose whether they will follow that decision or not. Moreover, it is also possible that making various expert systems using various philosophy of diagnosis. That is another merit of artificial intelligent system.

Orthodontics is the field that the expert system can be applied usefully.[19,21] The expert system constructed in this study showed high performances. In the near future, advanced computer technology could make it possible that automation of measuring of the diagnosis data.[35] Then an automatic process of the treatment planning might be possible.[20,36]

VI. Conclusions

We aimed to make an artificial intelligence model using neural network to diagnose extraction patterns. The subjects included 156 patients and were divided into the learning set and the test set. The learning set was used to make an artificial neural network model and the test set was used to evaluate the performance of this model. To make a well generalized model, the learning set was divided into the training set and the validation set. At the point where the error of validation set was minimal, the iterative learning with the training set was stopped.

The input data for the model consisted of 12 cephalometric variables and additional 6 indexes. The output data consisted of 3 bits to divide the extraction patterns. A 2-layer neural network including 1 hidden layer was selected for the machine learning. A backpropagation algorithm was applied to adjust the weighted values.

As a result of making models for the diagnosis of extractions using neural network machine learning, the success rates of the classifiers showed 93% for the diagnosis of extraction versus non-extraction and showed 84% for the detailed diagnosis of the extraction patterns in total.

This study suggests that artificial intelligence expert systems using neural network machine learning could be a new approach in orthodontics.

Table I. The subjects' sex, age, and other characteristics

Variables	n	Mean	SD
Age, y			
Female	94	25	7
Male	62	23	6
Type of extractions			
Non-extraction	62		
Ext44	20		
Ext45	36		
Ext55	25		
Ext40	13		
Type of learning			
Learning set	96		
Test set	60		
Total	156		

SD, standard deviation.

Table II. Descriptions for the six additional indexes

Index	Weighting	Criteria (mm)
Arch length discrepancy		
Spacing	0	ALD > 0
Normal	0.25	-1 < ALD ≤ 0
Mild crowding	0.5	-3 < ALD ≤ -1
Moderate crowding	0.75	-5 < ALD ≤ -3
Severe crowding	1	ALD ≤ -5
Molar key		
Class III key	0	
Super Class I key	0.25	
Class I key	0.5	
End-on key	0.75	
Class II key	1	
Large overjet		
Not severe	0	Overjet ≤ 5
Severe	1	Overjet > 5
Protrusion		
Concave profile	0	
Normal profile	0.25	
Mild protrusion	0.5	
Moderate protrusion	0.75	
Severe protrusion	1	
Chief complaint index for protrusion		
No protrusion in the CC	0	
Protrusion in the CC	1	

ALD, arch length discrepancy; CC, chief complaint.

Table III. Descriptions for the output data

Groups	Output data		
	Dx_ext	Dx_diff	Dx_more
Non-extraction	0		
Ext55	1	0	0
Ext44	1	0	1
Ext45	1	1	0
Ext40	1	1	1

Table IV. Decision-making success rates of the each classifier (%)

	Learning set		Test set	Total set
	Training set	Validation set		
Classifier_1	92 (59/64)	94 (30/32)	93 (56/60)	93 (145/156)
Classifier_2	88 (35/40)	100 (20/20)	85 (29/34)	89 (84/94)
Classifier_3	88 (21/24)	75 (6/8)	85 (11/13)	84 (38/45)
Classifier_4	95 (20/21)	100 (7/7)	95 (20/21)	96 (47/49)
Total	85 (82/96)		82 (49/60)	84 (131/156)

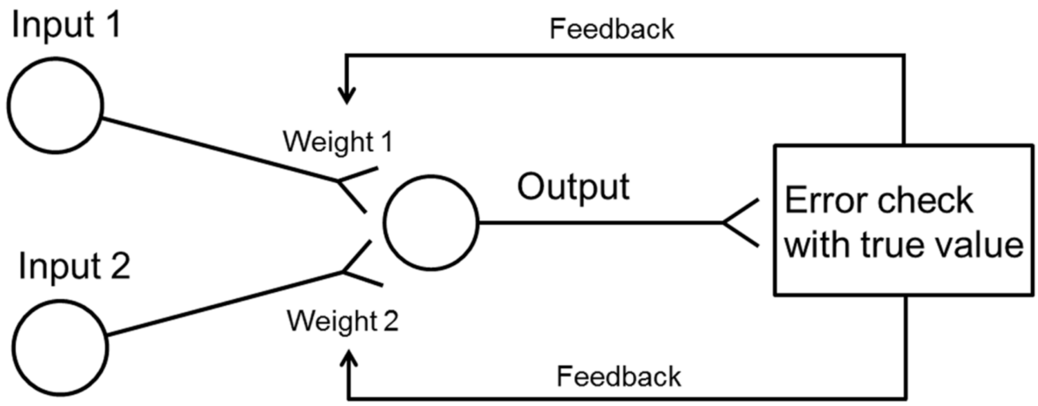

Fig 1. Schematic diagram of the neural network machine learning and weighting adjustment.

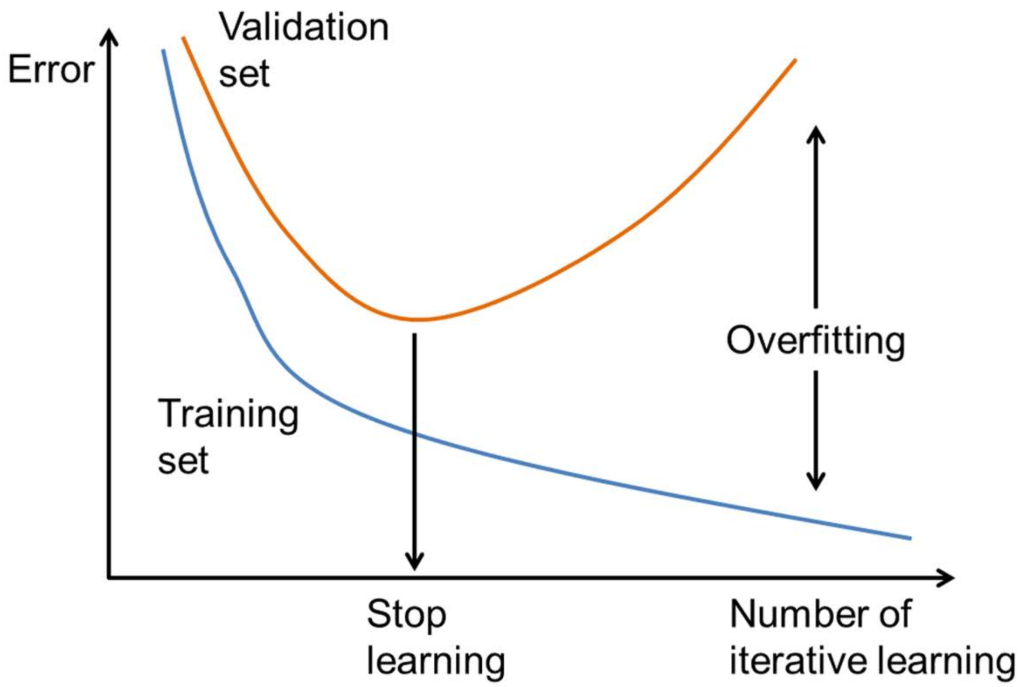

Fig 2. Learning curve of the training and validation sets.

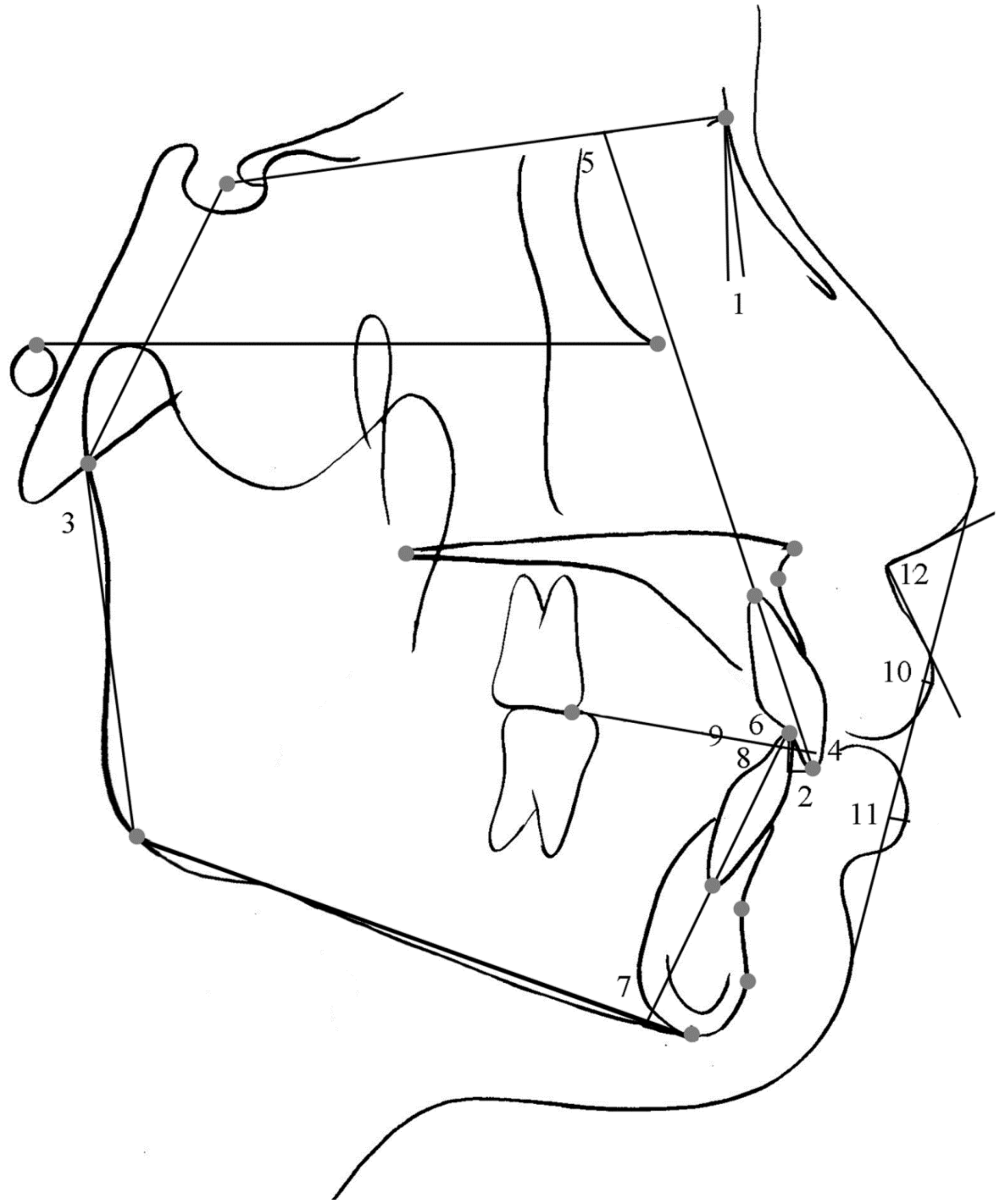

Fig 3. Linear and angular measurements used in this study.

1) ANB angle, 2) overjet, 3) Björk sum, 4) overbite, 5) maxillary central incisor to SN angle, 6) maxillary central incisor to occlusal plane angle, 7) IMPA, 8) mandibular central incisor to occlusal plane angle, 9) interincisal angle, 10) upper lip to E-line, 11) lower lip to E-line, and 12) nasolabial angle

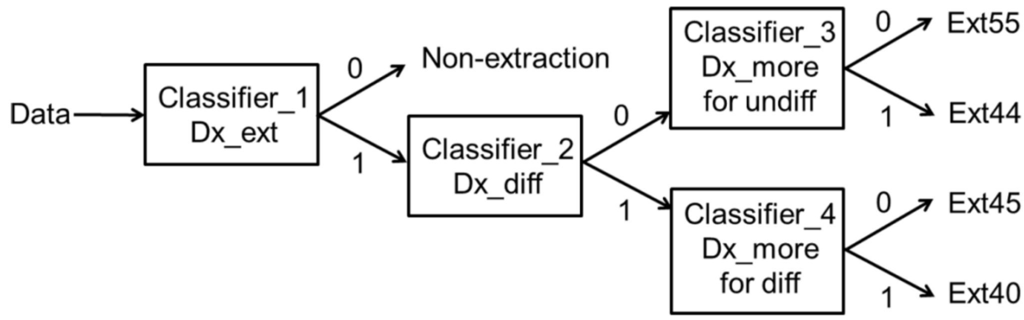

Fig 4. Schematic diagram of the stepwise learning used in this study.

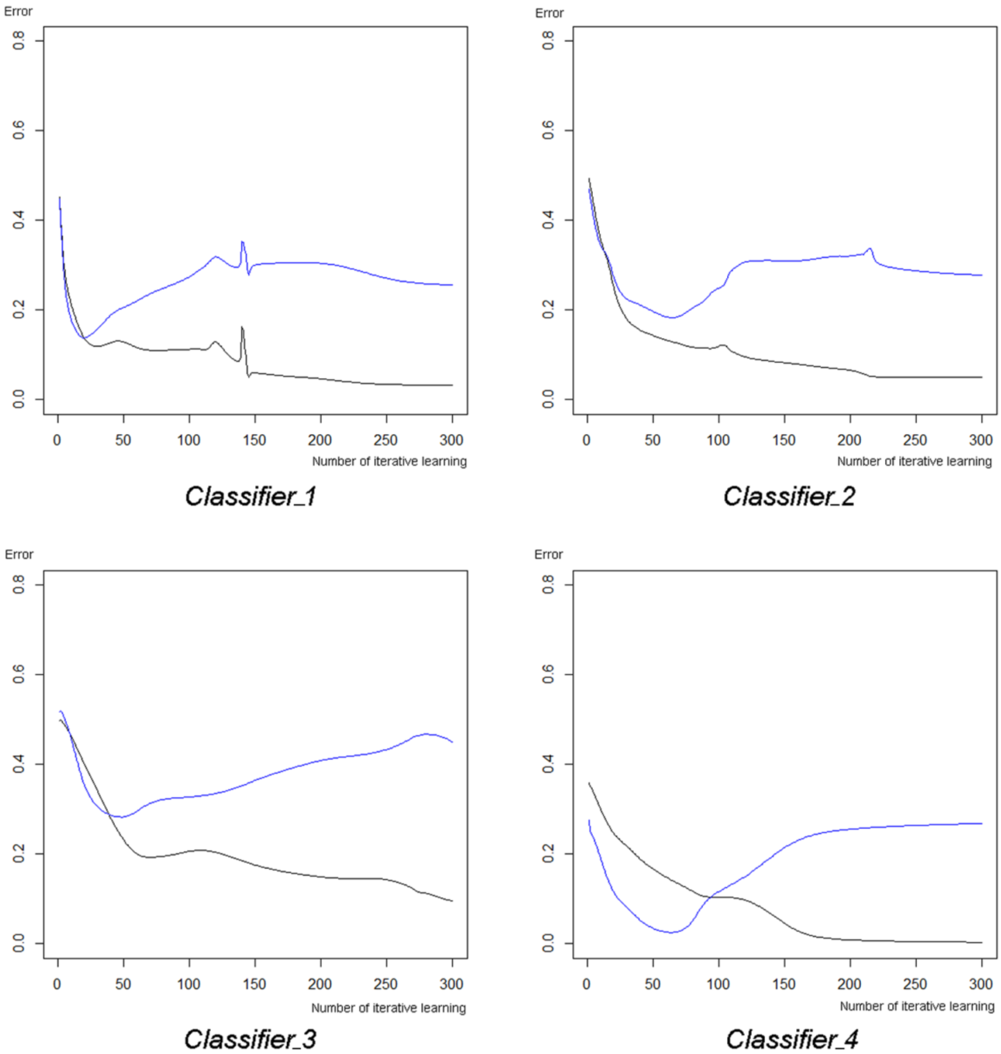

Fig 5. Learning curve (black) and validation curve (blue) of the each classifier.

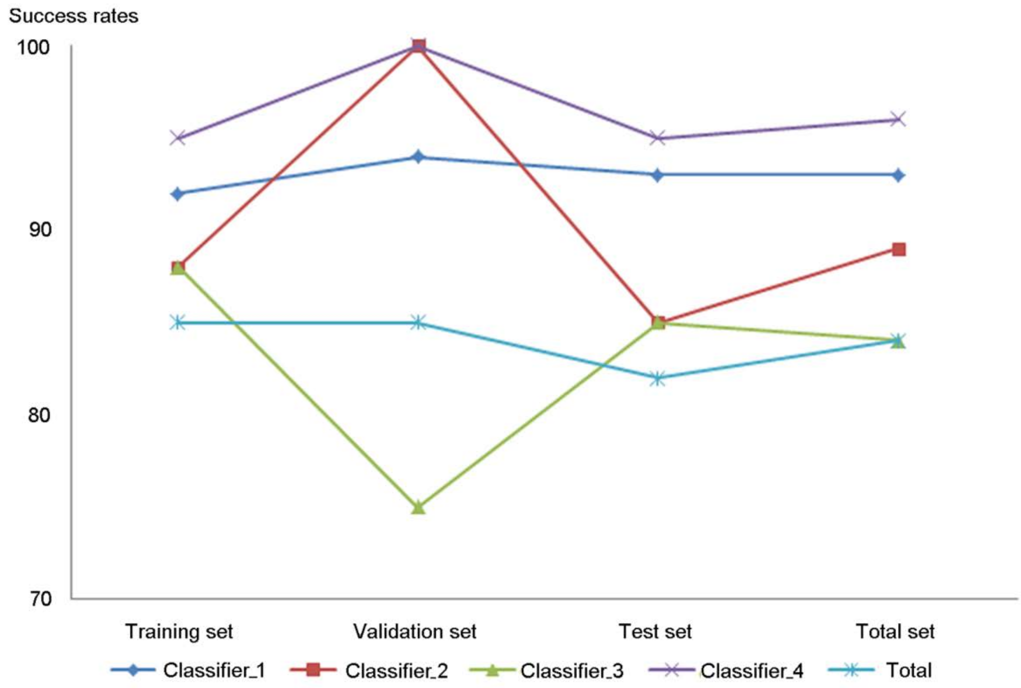

Fig 6. Success rates of the each classifier.

VII. References

1. Proffit WR, Fields HW, Sarver DM. Contemporary Orthodontics. 5th ed. St Louis, Mosby, 2013.

2. Ribarevski R, Vig P, Vig KD, Weyant R, O'Brien K. Consistency of orthodontic extraction decisions. Eur J Orthod 1996;18:77-80.

3. Dunbar AC, Bearn D, McIntyre G. The influence of using digital diagnostic information on orthodontic treatment planning - a pilot study. J Healthc Eng 2014;5:411-427.

4. Durao AR, Alqerban A, Ferreira AP, Jacobs R. Influence of lateral cephalometric radiography in orthodontic diagnosis and treatment planning. Angle Orthod 2015;85:206-210.

5. Rischen RJ, Breuning KH, Bronkhorst EM, Kuijpers-Jagtman AM. Records needed for orthodontic diagnosis and treatment planning: a systematic review. PLoS One 2013;8:e74186.

6. Devereux L, Moles D, Cunningham SJ, McKnight M. How important are lateral cephalometric radiographs in orthodontic treatment planning? Am J Orthod Dentofacial Orthop. 2011;139:e175-81.

7. Luke LS, Atchison KA, White SC. Consistency of patient classification in orthodontic diagnosis and treatment planning. Angle Orthod 1998;68:513-520.

8. Su MC, Chang HT. A new model of self-organizing neural networks and its application in data projection. IEEE Trans Neural Netw 2001;12:153-158.

9. Yamamoto Y, Nikiforuk PN. A new supervised learning algorithm for multilayered and interconnected neural networks. IEEE Trans Neural Netw 2000;11:36-46.

10. Mackin N, Sims-Williams JH, Stephens CD. Artificial intelligence in the dental surgery: an orthodontic expert system, a dental tool of tomorrow. Dent Update 1991;18:341-343.

11. Brickley MR, Shepherd JP, Armstrong RA. Neural networks: a new technique for development of decision support systems in dentistry. J Dent 1998;26:305-309.

12. Perfetti R, Ricci E. Analog neural network for support vector machine learning. IEEE Trans Neural Netw 2006;17:1085-1091.

13. Ying H, Lin F, MacArthur RD, Cohn JA, Barth-Jones DC, Ye H et al. A self-learning fuzzy discrete event system for HIV/AIDS treatment regimen selection. IEEE Trans Syst Man Cybern B Cybern 2007;37:966-979.

14. Sims J, Matthewman A, Brown D. An orthodontic expert system. Fuzzy Sets and Systems 1989;30:121-133.

15. Nieri M, Crescini A, Rotundo R, Baccetti T, Cortellini P, Pini Prato GP. Factors affecting the clinical approach to impacted maxillary canines: A Bayesian network analysis. Am J Orthod Dentofacial Orthop 2010;137:755-762.

16. Kennedy RL, Lee Y, Roy B, Reed CD, Lippman RP. Solving Data Mining Problems Through Pattern Recognition. 1st ed. Prentice Hall PTR,1997.

17. Xie X, Wang L, Wang A. Artificial neural network modeling for deciding if

extractions are necessary prior to orthodontic treatment. Angle Orthod 2010;80:262-266.

18. Brown ID, Adams SR, Stephens CD, Erritt SJ, Sims-Williams JH. The initial use of a computer-controlled expert system in the treatment planning of Class II division 1 malocclusion. Br J Orthod 1991;18:1-7.

19. Stephens C, Mackin N. The validation of an orthodontic expert system rule-base for fixed appliance treatment planning. Eur J Orthod 1998;20:569-578.

20. Noroozi H. Introduction of a new orthodontic treatment planning software; a fuzzy logic expert system. Int J Orthod Milwaukee 2006;17:25-29.

21. Hammond RM, Freer TJ. Application of a case-based expert system to orthodontic diagnosis and treatment planning. Aust Orthod J 1997;14:229-234.

22. Hicks EP, Kluemper GT. Heuristic reasoning and cognitive biases: Are they hindrances to judgments and decision making in orthodontics? Am J Orthod Dentofacial Orthop 2011;139:297-304.

23. Poon KC, Freer TJ. EICO-1: an orthodontist-maintained expert system in clinical orthodontics. Aust Orthod J 1999;15:219-228.

24. Lux CJ, Stellzig A, Volz D, Jager W, Richardson A, Komposch G. A neural network approach to the analysis and classification of human craniofacial growth. Growth Dev Aging 1998;62:95-106.

25. Yagi M, Ohno H, Takada K. Decision-making system for orthodontic treatment planning based on direct implementation of expertise knowledge. Conf

Proc IEEE Eng Med Biol Soc 2010;2010:2894-2897.

26. Takada K, Yagi M, Horiguchi E. Computational formulation of orthodontic tooth-extraction decisions. Part I: to extract or not to extract. Angle Orthod 2009;79:885-891.

27. Yu X, Liu B, Pei Y, Xu T. Evaluation of facial attractiveness for patients with malocclusion: a machine-learning technique employing Procrustes. Angle Orthod 2014;84:410-416.

28. Moghimi S, Talebi M, Parisay I. Design and implementation of a hybrid genetic algorithm and artificial neural network system for predicting the sizes of unerupted canines and premolars. Eur J Orthod 2012;34:480-486.

29. R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2013.

30. Fleiss JL, Levin B, Paik MC, Statistical Methods for Rates and Proportions, John Wiley & Sons Inc., Hoboken, New Jersey, 2003.

31. Han M, Snow PB, Brandt JM, Partin AW. Evaluation of artificial neural networks for the prediction of pathologic stage in prostate carcinoma. Cancer 2001;91:1661-1666.

32. Dayhoff JE, DeLeo JM. Artificial neural networks: opening the black box. Cancer 2001;91:1615-1635.

33. McEvoy FJ, Amigo JM. Using machine learning to classify image features from canine pelvic radiographs: evaluation of partial least squares discriminant

analysis and artificial neural network models. Vet Radiol Ultrasound 2013;54:122-126.

34. Bishara SE, Cummins DM, Jakobsen JR. The morphologic basis for the extraction decision in Class II, division 1 malocclusions: a comparative study. Am J Orthod Dentofacial Orthop 1995;107:129-135.

35. Forsyth DB, Davis DN. Assessment of an automated cephalometric analysis system. Eur J Orthod 1996;18:471-478.

36. Noroozi H. Orthodontic treatment planning software. Am J Orthod Dentofacial Orthop 2006;129:834-837.

국문초록

신경망 인공지능 의사결정 모델을 이용한 발치 진단의 새로운 방법 제안

서울대학교 대학원 치의과학과 치과교정학 전공
정 석 기

(지도교수 : 김 태 우)

1. 목적

이 연구의 목적은 신경망 기계학습을 이용하여 치료계획 수립의 과정 중 발치진단의 부분에 대한 임상가의 의사결정과정을 흉내낸 인공지능 모형을 만들고 그 효용성과 정확성을 평가함으로써 실제 임상에 활용될 수 있는지 여부를 타진해보고자 한다. 또한 이와 같은 과정을 통해 향후 교정진단의 다양한 분야에서 기계학습의 활용의 가능성을 알아보고자 한다.

2. 방법

샘플은 서울대학교 치과병원 치과교정과를 방문한 156명의 환자들을 대상으로 하였다. 교정경력 10년이상된 전문가의 결정을 바탕으로 다음의 5개의 치료계획그룹에 포함되는 샘플을 채택하였다: 비발치, 44/44 발치, 55/55발치, 44/55발치, 44/00발치. 이중 랜덤하게 배정된 96명을 가지

고 learning set을 구성하였으며 나머지 60명을 가지고 test set을 구성하였다. test set은 모형구성에 참여하지 않고 오로지 만들어진 모형평가에만 사용되었다. Overfitting을 방지하기 위해 Training set의 학습중 Validation set의 error를 최소로 하는 순간 학습을 멈추고 모형을 결정하였다. 이렇게 만들어진 모형들을 가지고 Test set에 적용함으로써 모형의 적절성 및 유효성을 평가하고 가장 적절하고 뛰어난 성능을 보이는 모형을 선택하였다.

Training은 3단계에 나누어 실시되었고 이를 통해 총 4개의 가장 성능이 뛰어난 모형이 채택되었다. 첫번째는 발치-비발치를 결정하는 모형이고 두번째는 발치케이스에서 대칭적 발치와 비대칭적 발치를 결정하는 모형을 만들었다. 세번째는 대칭적 발치케이스에서 발치심도를 결정하는 모형과 비대칭적 발치케이스에서 발치심도를 결정하는 모형 각각을 만들었다. 이렇게 만들어진 모형을 토대로 발치-비발치 결정 성공률, 대칭발치-비대칭발치 결정 성공률, 발치심도 결정 성공률을 계산하였으며 최종적으로 실제 진단과 인공지능모형을 통해 결정된 진단의 차이를 계산해 최종 진단 성공률을 계산하였다.

3. 결과

발치-비발치의 진단 성공률은 training set에서는 92% (59/64), validation set에서는 94% (30/32)로 나타났으며 test set에서는 93% (56/60), total 93% (145/156)으로 나타났다. 대칭-비대칭 발치 진단 성공률은 training set에서는 88% (35/40), validation set에서는 100% (20/20)로 나타났으며 test set에서는 85% (29/34), total 89% (84/94)으로 나타났다. 대칭 발치심도 진단 성공률은 training set에서는 88% (21/24), validation set에서는 75% (6/8), test set에서는 85% (11/13)로 나타났으며 total 84% (38/45)로 나타났다. 비대칭

발치심도 진단 성공률은 training set에서는 95% (20/21), validation set에서는 100% (7/7), test set에서는 95% (20/21)로 나타났으며 total 96% (47/49)로 나타났다. 4개의 의사결정 모델을 순차적으로 돌려 전체 data를 진단 검증한 결과는 학습 set에서는 85% (82/96), test set에서는 82% (49/60)으로 나타났으며 total 84% (131/156)으로 나타났다.

주요어: 기계학습, 발치진단, 신경망 모형
학번 : 2014-30728