Review Article

Applications of artificial intelligence and machine learning in orthodontics

Saeed N. Asiri[1,2], Larry P. Tadlock[1], Emet Schneiderman[3], Peter H. Buschang[1]

Departments of [1]Orthodontics, [3]Biomedical Sciences, College of Dentistry, Texas A&M University, Dallas, Texas, United States, [2]Department of Preventive Dental Sciences, College of Dentistry, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia.

*Corresponding author:
Saeed N. Asiri,
Department of Orthodontics,
College of Dentistry, Texas
A&M University, Dallas, Texas,
United States.

asiri@tamhsc.edu

Quick Response Code:
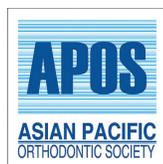
ABSTRACT

Over the past two decades, artificial intelligence (AI) and machine learning (ML) have undergone considerable development. There have been various applications in medicine and dentistry. Their application in orthodontics has progressed slowly, despite promising results. The available literature pertaining to the orthodontic applications of AI and ML has not been adequately synthesized and reviewed. This review article provides orthodontists with an overview of AI and ML, along with their applications. It describes state-of-the-art applications in the areas of orthodontic diagnosis, treatment planning, growth evaluations, and in the prediction of treatment outcomes. AI and ML are powerful tools that can be utilized to overcome some of the clinical problems that orthodontists face daily. With the availability of more data, better AI and ML systems should be expected to be developed that will help orthodontists practice more efficiently and improve the quality of care.

Keywords: Artificial intelligence, Machine learning, Artificial neural networks, Orthodontics, Dentistry

INTRODUCTION

Artificial intelligence (AI) is a subfield of computer science concerned with developing computers and programs that have the ability to perceive information, reason, and ultimately convert that information into intelligent actions.[1-3] AI as a science is very broad and encompasses various fields, including reasoning, natural language processing, planning, and machine learning (ML).[4] At present, ML is the most commonly used AI application in the medical and dental fields.

Work in AI started back in 1943,[5] but it was not until 1956 that the term "artificial intelligence" was first used during a conference held at Dartmouth College.[6] A few years later, the term "machine learning" was officially applied to a checkers-playing program, considered one of the first successful self-learning tools.[7] Drawing from other fields such as statistics, mathematics, physics, biology, neuroscience, and psychology,[8-11] AI and ML progressed quickly.

One of the most important aspects of any intelligent system is learning. Learning is the process of improving performance or behavior by practice and experience.[12] Similarly, ML is concerned with making machines and computers capable of learning from previous experiences, data, or examples. By utilizing a mixture of statistical and probabilistic tools, machines can learn from previous examples and improve their actions when new data are introduced. This could be in the form of predictions, identifying new patterns or classifying new data.[8] It is important to note that ML is not intended to mimic human behavior. Instead, it supplements human intelligence by

performing tasks that are beyond human capabilities.[13] This is what makes ML superior to the rule-based expert systems (ESs) that were used in the past.

ESs are considered among the earliest applications of AI. As the name implies, the knowledge of a specific field is transferred from humans to computers, allowing people to consult the computer.[14] In other words, ESs act as consultants that can process input information and provide solutions based on if-then rules. ESs have been used widely for diagnosis and treatment planning in medicine,[15] dentistry,[16] and orthodontics.[17] ESs also facilitate the transfer of knowledge to different people in different places. However, rule-based ESs are limited to information available at the time that the system is developed. Continuous updates are required to ensure that the information is correct and current. Due to the availability of more advanced technologies, such as ML, it is now possible to overcome the limitations associated with rule-based ES.

Most algorithms used in ML are also being used in data mining. The difference lies in the algorithm's goal. If the goal is to optimize decisions, then the algorithms are applied to large historical data sets to look for new patterns or relationships.[18,19] This process is called data mining. For example, data mining can help clinical practitioners find valuable information within existing patient records. Using this new information, practitioners can optimize future decisions, improve their daily practice, and increase the quality of care. On the other hand, if the goal is to make predictions, then ML should be applied. The clinical practitioner uses available data about a certain disease to train the machine to make predictions about the diagnosis or prognosis of patients that have never been seen before. Importantly, ML predictive models have proven to be more accurate than statistical models.[20] The aim of the present narrative review was two-fold: (1) To introduce the various types of ML and (2) show orthodontists how ML has been and is currently being applied. The literature was systematically searched using the MEDLINE (through PubMed) and ProQuest databases, covering both the published and unpublished literature reported in English. The studies covered are comprehensive with respect to orthodontic applications only.

TYPES OF ML

ML algorithms are divided into three main categories[10] [Figure 1] based on the nature of learning and the desired outcome of the algorithm:

Supervised learning

Supervised learning is mainly used for classification when the data are discrete (categorical) and for prediction (regression) if the data are continuous. It is supervised because it is based on a known outcome. With this type of learning, a model is built using a labeled set of training data (independent variables) and a known outcome (dependent) variable.[21] Since the final outcome is known, the system learns by receiving feedback signals that either confirm or reject its performance. If the algorithm encounters new input data, it will use the training data sets to link the new input data to the desired outcome. A very common example of supervised learning is e-mail spam detection, where the algorithm is trained to classify newly received emails as spam or not spam. For prediction, supervised learning can be used to predict the Graduate Record Examinations scores, for example, based on several independent variables that are related to the outcome variable, such as study time.

Unsupervised learning

This type of learning is mainly used to discover the structure of the data to find meaningful information. Clustering (sometimes called unsupervised classification) is the method used with this type of learning to explore the data and then organize it into groups based on similarities or relationships between variables.[21] Unlike supervised learning, the data are not labeled and the final outcome is not known. This type of learning allows marketers to develop programs that are specific to each group of customers after clustering them based on similar interests and features. The clusters could be based on sex, age group, or demographics.

Reinforcement learning

This type of learning is similar to supervised learning in that the system is provided with a feedback signal. However, the feedback signal does not provide the true value. Instead, it rewards the system based on its interaction with a dynamic environment (n.b. reinforcement learning is also known as the reward system). The system does not know anything about the behavior of the environment. By doing multiple exploratory trials and errors, the system learns and improves its future performance. An example of this type of learning is the chess engine. Depending on the situation (i.e., the environment), the machine decides on certain moves and will be rewarded by either winning or losing.[21]

MAJOR ML ALGORITHMS AND DENTISTRY

There are several ML algorithms that have been used in the dental fields. Depending on the goal, the type, and amount of data, different algorithms can be used. For example, if a practitioner wants to distinguish between patients who need treatment and those who do not, he/she probably would need to use a classification algorithm (e.g., support vector machine, naïve Bayes, etc.). However, if there are many variables and a large amount of data, an algorithm like neural networks is
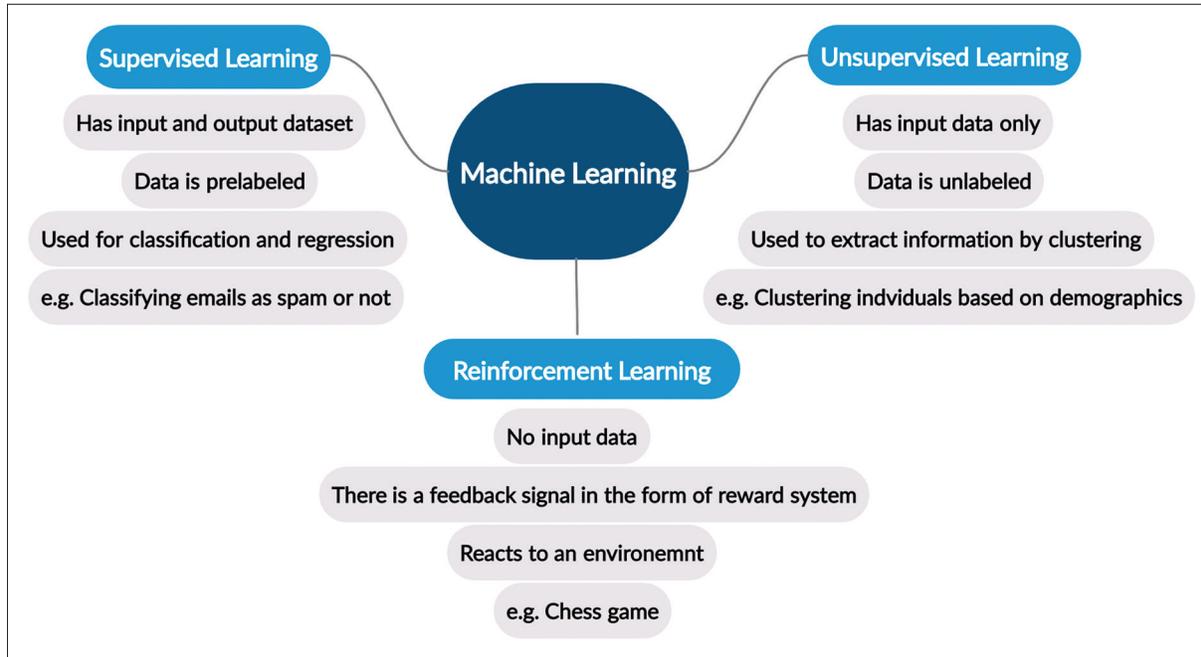
Figure 1: Types of machine learning.

better suited because it can handle noisy data and perform predictions even if the relationships between variables are non-linear.

Interestingly, almost all ML algorithms applied in orthodontics have used the supervised learning method [Table 1]. Most applications have sought to automate clinical procedures that perform or facilitate diagnosis and treatment planning. These applications require training with data that have a known and desired outcome.

AI AND ORTHODONTICS

Dentistry in general and orthodontic specifically has applied AI to solve many different problems. Early attempts to use AI in dentistry and orthodontics were in the form of knowledge-based ES. These systems were mainly aimed at helping non-specialist dentists develop diagnoses and treatment plans.[22-25] These ESs were useful in countries like England, where hospital-based orthodontists had long waiting lists and were seeing more patients than their counterparts in Europe and the US. Due to the decline in the incidence of caries that occurred at that time, dentists treated the more straightforward cases identified by the ES and referred the more complex cases to orthodontists. However, these systems were limited because they only had been introduced to simple cases (i.e., they could not function well with new cases not already stored in the system). At present, general dentists have more advanced ML systems available to them that can diagnose a broader range of orthodontic cases and determine treatment needs.[26] Several advanced systems have been developed to help orthodontists diagnose and treatment plan and evaluate treatment outcomes and growth.

ML FOR DIAGNOSIS AND ORTHODONTIC TREATMENT PLANNING

One of the dilemmas during treatment planning is deciding whether or not to extract, with substantial variability between orthodontists' decisions.[27] This has led to the development of several decision support systems that reduce the subjectivity of making decisions. Artificial neural networks (ANNs)[28-30] have been used to develop such systems, and they were shown to be successful at predicting the extraction decision 80%[28] of the time in one study and 93%[29,30] of the time in two other studies. Prediction of the detailed extraction patterns (i.e., which teeth needed to be extracted) was also shown to be possible 84%[29] of the time in one study and 83%[30] of the time in another study. Recently, a paper used ANN to identify anchorage requirements in cases that were determined by the system to need extractions and it was accurate 83% of the time.[30]

X-ray analysis, an integral part of diagnosis and treatment planning, has also benefited from ML. One of the most important applications of ML in orthodontics was the automation of landmark detections. A recent systematic review reported 5–15% better accuracy of landmark detection with ML than traditional methods.[31] ML was also used to automate diagnostics directly from cephalograms, including the sagittal relationships between the maxilla and mandible, as well as normal and abnormal posterior-anterior facial heights ratios, overbite, and overjet.[32]

Table 1: Summary of major machine learning algorithms applied in orthodontics.

Machine learning algorithm	Uses/applications	Pros	Cons
Decision trees	Used mainly for classification Applied in medical diagnosis[61] and manufacturing monitoring[62]	Simple and easy to understand even by non-experts[63] They are non-parametric and can handle both nominal and numeric input attributes[63] Can be used when data are missing, skewed, or have errors[64] Order of training instances is not important[65] Pruning reduces overfitting and improves prediction accuracy[65] Order of training has no effect on training[65]	Most algorithms require the target attribute to have only discrete values[63] They perform poorly when many complex interactions exist[63] Oversensitivity to the training set, irrelevant attributes and to noise[66]
Naïve Bayes	Used mainly for classification Applied in medicine[67,68] and dentistry[69,70] for decision support and risk assessment	Simple and easy to understand[71] Order of training has no effect on training[71] It is based on statistical modeling[71] Requires small amount of data for training[72] Fast and can deal with discrete and continuous attributes[72] Robust to outliers[73]	Accuracy is affected by redundant attributes and class frequency[71] Normal distribution is assumed for numeric attributes[71] Attributes are assumed to be conditionally independent[71]
Neural network	Used for classification and regression Applied in dentistry and medicine for diagnosis[37]	Boolean functions (AND, OR, and NOT) can be used with neural networks Can handle noisy inputs and allows changing input features during data collection[74] Successful with complex non-linear relationships between predicted variable and input data[74]	Overfitting is common especially with too many variables[75] Have limited ability to identify causal relationship[74] Require more computational resources[74]
Support vector machine	Used for classification and regression Applied in dentistry for classification of skeletal patterns[56]	Resistant to overfitting[10] Can model nonlinear functions[10] Can be used with non-linear relationships between predicted variable and input data	Training is slow Structure of algorithm is difficult to understand
Genetic algorithm	Used for search and optimization problems Applied in dentistry and medicine mainly for prediction	Simple algorithm and easy to apply[76] Always try to find the best solution	Not efficient for finding the best solution There are complications in representing training and output data
Fuzzy logic	Concerned with finding the truth by approximate modes of reasoning rather than exact reasoning[77] Used to deal with imprecision and uncertainty present in many fields including medicine[78]	Mimics human thinking and can be written in a form similar to natural language[79] Allows for the degree of belonging to either 0 or 1, with 1 representing complete membership and 0 for non-membership Can use both numerical variables and linguistic variables[80]	Requires a lot of data and expertise to develop[81] Analysis is difficult because fuzzy outputs can be interpreted in different ways[81]

Automation of X-rays analysis has also been extended to hand and wrist radiographs for estimating skeletal age. Determining the growth status of patients is essential for deciding whether or not to utilize growth during treatment.[33] A ML system

applied to a sample of 360 images showed an average difference of 0.39 years between its estimate and skeletal age estimated by two expert radiologists.[34] Another study, using a larger sample of 1100 images, reported an average difference of 0.60 years when compared to two experienced radiologists.[35] One study comparing the performance of different algorithms to estimate skeletal age reported a root-mean-square error (RMSE) of 0.24 years with ANN and 0.25 years with a genetic algorithm when compared to traditional estimates of skeletal age.[36]

Taking panoramic radiographs make orthodontists legally liable if they overlook diagnosing a lesion or a tumor. This has led to the development of an automated neural network system that can correctly diagnose ameloblastomas and keratocystic odontogenic tumors from panoramic radiographs 83.0% of the time.[37] Five oral and maxillofacial surgeons who examined the same radiographs correctly diagnosed the problems 82.9% of the time. The difference lies in the time needed for diagnosis. The ML system required an average of 38 s, while the surgeons needed 23.1 min for each diagnosis. Another system was developed that successfully predicted odontogenic cysts, dentigerous cysts, osteomyelitis, periapical cysts, and ameloblastomas 90.6%, 90.9%, 99.4%, 89.6%, and 100% of the time, respectively.[38] At present, more and more orthodontists are using cone-beam computed tomography, which has led to the development of an automated system using the support vector machine to correctly diagnose periapical cysts and keratocystic odontogenic tumors 100% of the time.[39] Neural networks were used to estimate patients' dental ages from panoramic radiographs.[40] Its RMSE was 0.9 for girls and 1.1 for boys, while traditional regression had an RMSE of 1.3 and 1.4 for girls and boys, respectively.[40]

Panoramic and lateral cephalometric X-rays have also been used to predict maxillary canine impactions based on angular and linear measures.[41] The highest prediction accuracy was obtained with a random forest algorithm, which correctly predicted the actual eruption status of canines 88.3% of the time.

One of the challenges for less experienced orthodontists is the selection of the appropriate treatment modality and appliance, including headgears. To address this, a system was developed to help orthodontists select the headgears that should be used.[42] Compared to the selections made by eight expert orthodontists, the system correctly identified the appropriate headgears 95.6% of the time. Recently, decision support systems were developed to determine the geometry of orthodontic springs used to close extraction spaces[43] and to determine the forces needed to align teeth,[44] but neither system has been applied clinically.

Another orthodontic challenge during treatment planning is predicting the size of unerupted teeth. To address this, a hybrid system using ANN and genetic algorithms was used to predict canine and premolar sizes.[45] Its maximum error was 2.4 mm for the mandibular and 1.6 mm for the maxillary teeth. The errors were often half as large as the error produced with linear regression prediction models.

ML AND TREATMENT OUTCOMES

One of the more useful applications of AI in orthodontics is the prediction of soft tissues treatment outcomes. Recently, ANN was used to predict the change in lip curvature after orthodontic treatment with or without extractions.[46] Its prediction of change and the actual change that occurred differed by 29.6% and 7% for the upper and lower lips, respectively. Both predictions were much better than those based on linear regression.

The topic of beauty is controversial because it is subjective and affected by factors such as age, sex, and ethnic backgrounds. Using ANN, facial attractiveness was quantified on a scale from 0 to 100 (0 extremely unattractive and 100 extremely attractive) before and after orthognathic surgery.[47] The difference between the pre- and post-surgery scores was shown to be statistically significant, with facial attractiveness improving 74.7%.

Predictions of treatment outcomes in Class II and Class III patients have also been reported. Using ANN, predictive models were developed to predict the post-treatment peer assessment rating (PAR) index in Class II patients based on their pre-treatment PAR index.[48] The neural network model used in this system was able to correctly predict the final PAR score 94.0% of the time; linear regression was correct only 82.0% of the time. A system has also been developed to predict outcomes in untreated Class III patients.[49] Unsupervised learning was used to cluster patients as hypermandibular, hyperdivergent, or balanced based on cephalometric variables. The system was then applied to a treated sample, where it showed that all of the unsuccessful cases belonged to either the hypermandibular or the hyperdivergent cluster. Another system was able to correctly predict the prognosis of Class III treatment 97.2% of the time, which was slightly better than 92.1% reported for discriminant analysis.[50]

ML AND GROWTH PATTERNS

Several methods have been introduced to help orthodontists classify their patients' growth patterns.[51-53] In 1998, an ANN was used to classify the growth of 43 untreated children based on size and shape changes.[54] However, the system was not validated on an external sample. A recent study used cephalometric variables to classify patients' craniofacial growth as either normal or abnormal.[55] It showed that support vector machines could correctly classify abnormal growth patterns 99.8% of the time. Another study using support vector machines to classify normal or abnormal

skeletal patterns based on craniofacial measures was correct only 74.5% of the time.[56]

Classification of Class III growth patterns has also been performed. Based on longitudinal data of untreated Class III subjects, who were classified as either good or bad growers based on the changes in their sagittal relationships, a classification tree had a significantly lower rate of misclassification (12.0%) than discriminant analysis (40.7%), both of which were based based on the same 11 cephalometric variables.[57] When the system was tested on new data, it was able to successfully identify good and bad growth patterns 64.0% of the time.

CONCLUSIONS

AI and ML systems applied in orthodontics provide promising tools that can improve clinical practice. These clinical decision support systems can help orthodontists practice more efficiently, reduce variability, and eliminate subjectivity.[58] The accuracy of most systems presently available is considered good to excellent ranging from approximately 64% to 97%. The accuracy at the lower end of this range should be expected to improve in the future as sample sizes increase and more information becomes available. Most of the systems were developed using restricted samples that reduce their generalizability. For example, patients were often excluded because they needed surgery or had missing teeth, unusual extraction patterns, or asymmetries. Future studies are needed to build predictive models that include different types of patients. Algorithms should also be expected to improve, making it possible to handle more complex data such as images. Systems based on images require more time, experience, and training data than systems based on discrete or continuous data values. This is especially important in the era of digital dentistry, where all patient's records such as dental models, X-rays, and facial photos are stored in computers in the form of digital images.

It is important to note that AI models are limited and have drawbacks. They should be used only after careful considerations. Like any statistical model, the ML algorithms are based on assumptions and have limitations. If used incorrectly, they can give misleading information. In addition, the quality of data is very important.[59] Data with a lot of noise, missing information, and more variables than observations can result in poor models. Moreover, the phenomena called overfitting occurs when a model is trained too many times on too few observations.[60] Such models perform poorly when introduced to new data. Keeping that in mind, orthodontists should understand that these AI models are meant to assist with the clinical judgment and not to substitute for the knowledge and expertise of humans.

Declaration of patient consent

Patient's consent not required as patient's identity is not disclosed or compromised.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Haugeland J. Artificial Intelligence: The Very Idea. Cambridge, MA: Massachusetts Insitute of Technology; 1985.
2. Morris CG. Academic Press Dictionary of Science Technology. San Diego: Academic Press; 1996.
3. Luger GF. Artificial Intelligence: Structures and Strategies for Complex Problem Solving. Harlow: Pearson Education; 2005.
4. Mueller JP, Massaron L. Machine Learning for Dummies. New York: John Wiley and Sons; 2016.
5. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 1943;5:115-33.
6. Negnevitsky M. Artificial Intelligence: A Guide to Intelligent Systems. Canada: Pearson Education; 2005.
7. Samuel AL. Some studies in machine learning using the game of checkers. IBM J Res Dev 2000;44:206-26.
8. Marsland S. Machine Learning: An Algorithmic Perspective. 2nd ed. London: Chapman and Hall/CRC; 2014.
9. Raina CK. A review on machine learning techniques. Int J Recent Innov Trends Comput Commun 2016;4:395-9.
10. Russell SJ, Norvig P. Artificial Intelligence: A Modern Approach. 3rd ed. Upper Saddle River, NJ: Pearson Education; 2011.
11. Coppin B. Artificial Intelligence Illuminated. Burlington: Jones and Bartlett Publishers; 2004.
12. Lachman SJ. Learning is a process: Toward an improved definition of learning. J Psychol 1997;13:477-480.
13. Shalev-Shwartz S, Ben-David S. Understanding Machine Learning: From Theory to Algorithms. Cambridge: Cambridge University Press; 2014.
14. Liang TP, Turban E, Aronson JE. Decision Support Systems and Intelligent Systems. 7th ed. Upper Saddle River, NJ: Prentice-Hall; 2004.
15. Borgohain R, Sanyal S. Rule based expert system for diagnosis of neuromuscular disorders. Int J Adv Netw Appl 2012;4:1509-13.
16. Stheeman S, Van der Stelt P, Mileman P. Expert systems in dentistry. Past performance future prospects. J Dent 1992;20:68-73.
17. Grant PD, Freer TJ. Expert systems in orthodontic clinical diagnosis. Aust Ortho J 1993;13:32.
18. Sumathi S, Sivanandam S. Introduction to data mining principles. In: Introduction to Data Mining and its Applications. Studies in Computational Intelligence. Berlin, Heidelberg: Springer; 2006. p. 1-20.
19. Mitchell TM. Machine learning and data mining. Commun ACM 1999;42:1-13.
20. Breiman L. Statistical modeling: The two cultures (with comments and a rejoinder by the author). Stat Sci

2001;16:199-231.

21. Raschka S. Python Machine Learning. Birmingham: Packt Publishing Ltd.; 2015.

22. Sims-Williams J, Brown I, Matthewman A, Stephens C. A computer-controlled expert system for orthodontic advice. Br Dent J 1987;163:161-6.

23. Williams JS, Matthewman A, Brown D. An orthodontic expert system. Fuzzy Sets Syst 1989;30:121-33.

24. Brown I, Adams S, Stephens C, Erritt S, Sims-Williams J. The initial use of a computer-controlled expert system in the treatment planning of class II division 1 malocclusion. Br Dent J 1991;18:1-7.

25. Hammond R, Freer TJ. Application of a case-based expert system to orthodontic diagnosis and treatment planning: A review of the literature. Aust Ortho J 1996;14:150-3.

26. Thanathornwong B. Bayesian-based decision support system for assessing the needs for orthodontic treatment. Healthc Inform Res 2018;24:22-8.

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

28. Xie X, Wang L, Wang A. Artificial neural network modeling for deciding if extractions are necessary prior to orthodontic treatment. Angle Orthod 2010;80:262-6.

29. Jung SK, Kim TW. New approach for the diagnosis of extractions with neural network machine learning. Am J Orthod Dentofac Orthop 2016;149:127-33.

30. Li P, Kong D, Tang T, Su D, Yang P, Wang H, *et al*. Orthodontic treatment planning based on artificial neural networks. Sci Rep 2019;9:2037.

31. Leonardi R, Giordano D, Maiorana F, Spampinato C. Automatic cephalometric analysis: A systematic review. Angle Orthod 2008;78:145-51.

32. Arik SÖ, Ibragimov B, Xing L. Fully automated quantitative cephalometry using convolutional neural networks. J Med Imaging 2017;4:014501.

33. Buschang PH, Roldan SI, Tadlock LP. Guidelines for assessing the growth and development of orthodontic patients. Semin Orthod 2017;23:321-35.

34. Giordano D, Kavasidis I, Spampinato C. Modeling skeletal bone development with hidden Markov models. Comput Methods Programs Biomed 2016;124:138-47.

35. Kashif M, Deserno TM, Haak D, Jonas S. Feature description with SIFT, SURF, BRIEF, BRISK, or FREAK? A general question answered for bone age assessment. Comput Biol Med 2016;68:67-75.

36. Mansourvar M, Shamshirband S, Raj RG, Gunalan R, Mazinani I. An automated system for skeletal maturity assessment by extreme learning machines. PLoS One 2015;10:e0138493.

37. Poedjiastoeti W, Suebnukarn S. Application of convolutional neural network in the diagnosis of jaw tumors. Healthc Inf Res 2018;24:236-41.

38. Yong TH, Lee SJ, Yi WJ. Odontogenic Cysts and Tumors Detection in Panoramic Radiographs using Deep Convolutional Neural Network (DCNN). In: Conference Proceedings of Internationa Conference on Medical Imaging with Deep Learning MIDL; 2019. p. 1-4.

39. Yilmaz E, Kayikcioglu T, Kayipmaz S. Computer-aided diagnosis of periapical cyst and keratocystic odontogenic tumor on cone beam computed tomography. Comput Methods Programs Biomed 2017;146:91-100.

40. Velemínská J, Pílný A, Čepek M, Koťová M, Kubelková R. Dental age estimation and different predictive ability of various tooth types in the Czech population: Data mining methods. Anthropol Anz 2013;70:331-45.

41. Laurenziello M, Montaruli G, Gallo C, Tepedino M, Guida L, Perillo L, *et al*. Determinants of maxillary canine impaction: Retrospective clinical and radiographic study. J Clin Exp Dent 2017;9:e1304-9.

42. Akçam MO, Takada K. Fuzzy modelling for selecting headgear types. Eur J Orthod 2002;24:99-106.

43. Lim D, Ong YS, Setiawan R, Idris M, editors. Classifier-assisted Constrained Evolutionary Optimization for Automated Geometry Selection of Orthodontic Retraction Spring. In: Conference Proceedings IEEE Congress on Evolutionary Computation; 2010. p. 1-8.

44. Omran LN, Ezzat KA, Hassanien AE. Decision Support System for Determination of Forces Applied in Orthodontic Based on Fuzzy Logic. In: The International Conference on Advanced Machine Learning Technologies and Applications. AMLTA. Advances in Intelligent Systems and Computing. Vol. 723. Cham, Switzerland: Springer; 2018.

45. 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 2011;34:480-6.

46. Nanda SB, Kalha AS, Jena AK, Bhatia V, Mishra S. Artificial neural network (ANN) modeling and analysis for the prediction of change in the lip curvature following extraction and non-extraction orthodontic treatment. J Dent Spec 2015;3:217-9.

47. Patcas R, Bernini D, Volokitin A, Agustsson E, Rothe R, Timofte R. Applying artificial intelligence to assess the impact of orthognathic treatment on facial attractiveness and estimated age. Int J Oral Maxillofac Surg 2018;48:77-83.

48. Zarei A, El-Sharkawi M, Hairfield M, King G. An Intelligent System for Prediction of Orthodontic Treatment Outcome. In: Proceedings of the IEEE International Joint Conference on Neural Network; 2006. p. 2702-6.

49. Auconi P, Scazzocchio M, Cozza P, McNamara JA Jr., Franchi L. Prediction of class III treatment outcomes through orthodontic data mining. Eur J Orthod 2015;37:257-67.

50. Kim BM, Kang BY, Kim HG, Baek SH. Prognosis prediction for class III malocclusion treatment by feature wrapping method. Angle Orthod 2009;79:683-91.

51. Brodie AG. Behavior of normal and abnormal facial growth patterns. Am J Orthod Dent Orthop 1941;27:633-47.

52. Broadbent BH. The face of the normal child. Angle Orthod 1937;7:183-208.

53. Ricketts RM. A principle of arcial growth of the mandible. Angle Orthod 1972;42:368-86.

54. Lux C, Stellzig A, Volz D, Jäger 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.

55. Lakkshmanan A, Shri AA, Aruna E. Pattern Classification for Finding Facial Growth Abnormalities. In: Proceedings of the International Conference on Computational Intelligence and Computing Research; 2013. p. 1-5.

56. Niño-Sandoval TC, Perez SV, González FA, Jaque RA, Infante-Contreras C. An automatic method for skeletal patterns classification using craniomaxillary variables on a Colombian population. Forensic Sci Int 2016;159:e1-6.

57. Auconi P, Scazzocchio M, Caldarelli G, Nieri M, McNamara JA, Franchi L. Understanding interactions among cephalometrics variables during growth in untreated Class III subjects. Eur J Orthod 2017;39:395-401.

58. Khanna S. Artificial intelligence: Contemporary applications and future compass. Int Dent J 2010;60:269-72.

59. Botsis T, Hartvigsen G, Chen F, Weng C. Secondary use of EHR: Data quality issues and informatics opportunities. Summits Transl Bioinform 2010;2010:1-5.

60. Rodvold DM, McLeod DG, Brandt JM, Snow PB, Murphy GP. Introduction to artificial neural networks for physicians: Taking the lid off the black box. Prostate 2001;46:39-44.

61. Decaestecker C, Remmelink M, Salmon I, Camby I, Goldschmidt D, Petein M, *et al*. Methodological aspects of using decision trees to characterise leiomyomatous tumors. Cytometry 1996;24:83-92.

62. Du R, Elbestawi M, Wu S. Automated monitoring of manufacturing processes, part 1: Monitoring methods. J Eng Ind 1995;117:121-32.

63. Rokach L, Maimon O. Decision trees. In: Maimon O, Rokach L, editors. Data Mining and Knowledge Discovery Handbook. Berlin, Heidelberg: Springer; 2005. p. 165-92.

64. Song YY, Ying L. Decision tree methods: Applications for classification and prediction. Shanghai Arch Psychiatry 2015;27:130-5.

65. Quinlan JR. Induction of decision trees. Mach Learn 1986;1:81-106.

66. Quinlan JR. C4. 5: Programs for Machine Learning. Amsterdam: Elsevier; 1993.

67. Pattekari SA, Parveen A. Prediction system for heart disease using Naïve Bayes. Int J Adv Comput Math Sci 2012;3:290-4.

68. Subbalakshmi G, Ramesh K, Rao MC. Decision support in heart disease prediction system using naive bayes. Ind J Comput Sci Eng 2011;2:170-6.

69. Chattopadhyay S, Davis RM, Menezes DD, Singh G, Acharya RU, Tamura T. Application of Bayesian classifier for the diagnosis of dental pain. J Med Syst 2012;36:1425-39.

70. Suebnukarn S, Rungcharoenporn N, Sangsuratham S. A Bayesian decision support model for assessment of endodontic treatment outcome. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2008;106:e48-58.

71. Langley P, Iba W, Thompson K, editors. An analysis of Bayesian classifiers. In: Proceedings of the 10th National Conference on Artificial Intelligence; 1992. p. 1-14.

72. Bhargavi P, Jyothi S. Applying naive bayes data mining technique for classification of agricultural land soils. Int J Comput Sci Netw Sec 2009;9:117-22.

73. Boyles S, Fajardo D, Waller ST, editors. A Naive Bayesian Classifier for Incident Duration Prediction. In: The 86th Annual Meeting of the Transportation Research Board; 2007. p. 1-10.

74. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol 1996;49:1225-31.

75. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323:533-6.

76. Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. Cancer Inform 2007;2:59-77.

77. Zadeh LA. Fuzzy logic. Computer 1988;21:83-93.

78. Torres A, Nieto JJ. Fuzzy logic in medicine and bioinformatics. J Biomed Biotechnol 2006;2006:1-7.

79. Warren J, Beliakov G, Van Der Zwaag B. Fuzzy Logic in Clinical Practice Decision Support Systems. In: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences; 2000. p. 1-10.

80. Zadeh LA. Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans Syst Man Cybern Syst 1973;1:28-44.

81. Godil SS, Shamim MS, Enam SA, Qidwai U. Fuzzy logic: A "simple" solution for complexities in neurosciences? Surg Neurol Int 2011;2:24.