



Review Article

## The role of AI and machine learning in contemporary orthodontics

Jean-Marc Retrouvey

Department of Orthodontics, School of Dentistry, University of Missouri, Kansas City, Missouri, United States.



**\*Corresponding author:**

Jean-Marc Retrouvey,  
Department of Orthodontics,  
School of Dentistry, University  
of Missouri, Kansas City,  
Missouri, United States.

[jean-marc.retrouvey@umkc.edu](mailto:jean-marc.retrouvey@umkc.edu)

Received : 04 September 2020

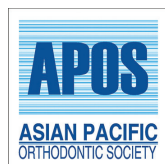
Accepted : 19 November 2020

Published : 10 April 2021

**DOI**

10.25259/APOS\_148\_2020

**Quick Response Code:**



### ABSTRACT

In the past 20 years, the orthodontic profession has adopted digital technologies such as computer-assisted tooth movement, automated staged dental aligner production, and 3D simulations. Until recently, the use of artificial intelligence (AI) was limited to narrow intelligence and supervised “learning” such as automated cephalometric point recognition, segmentation of teeth from 3D files, and staging of orthodontic treatment. The next step will be to create neural networks based on general intelligence (the human intelligence is considered general intelligence), where the network using powerful computers and complex algorithms will “learn” orthodontic diagnosis and treatment planning to suggest the most appropriate treatment plan for optimized treatments and more predictable outcomes. The objectives of this paper are to describe the state of the art in AI and orthodontics and explore potential avenues for future developments and applications.

**Keywords:** Artificial intelligence, Machine learning, Orthodontics, Diagnosis, Treatment planning

### INTRODUCTION

Since Invisalign™ entered the market in 1998 with its cloud based software called Clincheck,<sup>[1]</sup> the world of orthodontics has evolved dramatically. Align developed a novel concept where the orthodontists submit their patient’s records or dataset in return for a suggested treatment plan and subsequently the production of a set of aligners that move the teeth in predetermined position.<sup>[2]</sup>

The public embraced the concept, and soon, most of the orthodontic profession joined the growing number of providers submitting their cases to the company in return for a set of aligners. The process was pushed one step further with the submission of sequential records used for “refinement” and final records used for retention.<sup>[3]</sup>

This process is a unilateral data sharing<sup>[4]</sup> “association.” All patient data gathered by the orthodontists and dental professionals alike are submitted without any covenant of usage to the align company allowed this company to build an extraordinary and unique dataset of orthodontic treatment outcomes never before gathered by any orthodontic organization.<sup>[5]</sup>

Orthodontics is an art and a science based primarily on the experience and bias of the clinician.<sup>[6,7]</sup> Each malocclusion is unique, and it is not possible to totally and predictably correlate the different patterns expressed by the stomatognathic system.<sup>[8]</sup> Orthodontists usually do not share information among themselves except for the before and after records of patients at conferences with a strong bias toward publishing mainly positive outcomes.

This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-Share Alike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

©2021 Published by Scientific Scholar on behalf of APOS Trends in Orthodontics

Very little analytic randomized data are obtained from which even less is shared among practitioners. Clinical orthodontists treat one patient at a time and are not equipped to gather or share large dataset of treatments outcome to draw inferences from multiple and asynchronous clinical observations.<sup>[9]</sup> Orthodontists oftentimes rely on a limited amount of approaches that give them the perceived maximum efficiency. This approach based on clinical experience may frequently result in delayed treatment completion time and potentially less than optimal outcomes.<sup>[10]</sup>

Artificial intelligence (AI)<sup>[11]</sup> offers “a way to get sharper prediction from data”<sup>[12,13]</sup> by simultaneously analyzing all the different variables present in a malocclusion. This capacity offers the potential to assist the practitioner to obtain the most favorable outcome when treating a malocclusion.<sup>[14]</sup>

## CONVENTIONAL DIAGNOSTIC PROCEDURE

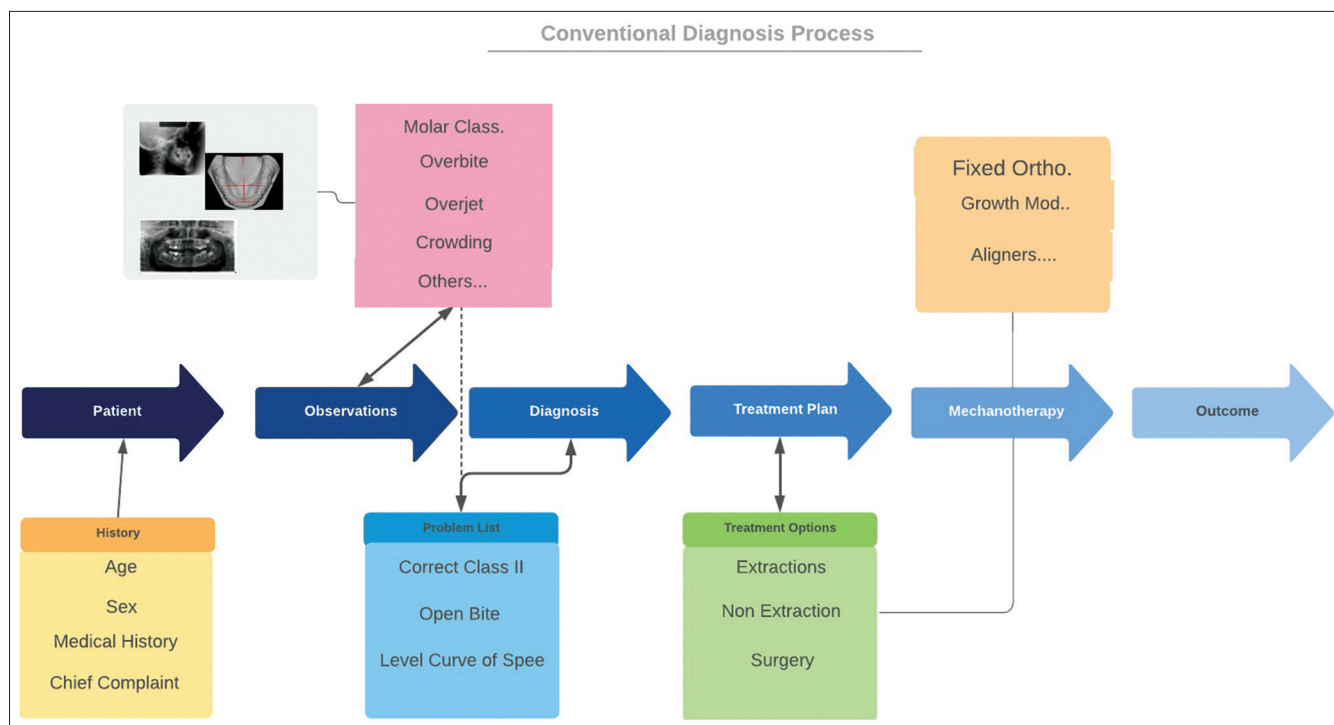
Studying and diagnosing a malocclusion present many challenges and bring uncertainty to the outcome of treatment due to the large number of variables present in the analysis.<sup>[9]</sup> The orthodontist must compute mentally all the parameters to recognize patterns based on experience and adopt the most logical or probable approach to solve the problem presented.<sup>[15,16]</sup> To simplify the analytic process, many orthodontists tend to simplify the analytic process tend to adopt a mechanical scheme over an in-depth diagnostic approach as it is perceived that experience alone will be enough to assess the probability of success.<sup>[9]</sup>

This process, called the “feedforward approach,”<sup>[17]</sup> does not necessarily involve complex feedback mechanisms to improve on previous diagnostics and outcome analyses.<sup>[13]</sup> [Figure 1] explains this principle as an accepted but potentially inefficient pattern to treat patients. This experience-based approach has been advocated by many clinicians as it is simple but lacks the means to re-evaluate and learn from positive and/or negative outcomes.<sup>[18]</sup>

The “Premeditatio Malorum” looks at a situation and explores all the options where the situation may go wrong to avoid such negative condition. Wouldn't this approach be a welcomed addition to the feed-forward diagnosis process?<sup>[19]</sup> Finding potential negative outcomes and then going backward to avoid creating errors in the diagnosis procedure would add more insight than a linear and unidirectional diagnosis process.<sup>[20]</sup>

## BIG DATA, AI, AND MACHINE LEARNING

Machine learning algorithms have existed for a very long time.<sup>[21]</sup> This field of computer science was initially derived from the field of statistics and then branched out into its own discipline.<sup>[22]</sup> The latest evolution in this field of research came with advancements in computing power with the introduction of hypercomputers and the use of graphic processing units instead of central processing units. These advances allowed for the acquisition of large amounts of data, known as “Big Data.” Big data came to life in the early 2000 following the publication of a paper by Laney describing



**Figure 1:** Forward diagnostic process using the “feed forward” approach without feedback.

the 3Vs (volume, velocity, and variety). \*Laney Datafication or the constant collection of data is creating an increasingly large amount of data that drives innovation.<sup>[23,24]</sup> In recent years, the greatest advancement in the field of machine learning was the introduction of deep learning.<sup>[25]</sup> AI and deep learning are the new buzzwords in dentistry and they rely on big data.<sup>[26]</sup>

Computer hardware has evolved dramatically, and now, large amount of data can be acquired and stored very efficiently. Rapid improvements in computing power allowed to train more sophisticated models to accomplish even more or evermore complex tasks, that is, convolutional neural networks for image detection.<sup>[27]</sup> Data are the key ingredient of machine learning<sup>[28]</sup> as Modern software programs, sophisticated statistical analyzes, and algorithms require a tremendous amount of data to be able to predict treatment outcomes and analyze the shortcoming of past-approaches. Neural networks use multiple algorithms to attempt to mimic the human brain and based on Bayesian probabilistic graphical models.<sup>[29]</sup> They require “large datasets to answer queries (questions) regarding a set of standard variables.”<sup>[30]</sup>

### What is a neural network and how does it eventually apply to orthodontic diagnosis?

Neural networks consist of nodes loosely referred as neurons.<sup>[31]</sup> Each node corresponds to a random variable. There are as many nodes as there are columns into a conventional data table. These nodes require inputs such as amount of overbite, overjet, and crowding which are initially randomly weighted (0 to 1%) and linked to nodes which calculate the correlation in the input layer. To further refine, the calculations and the aptitude of the network to “learn” from the data provided, “hidden layers” are added and create

deep learning neural networks to strengthen the probabilities that the neural network predictions will be accurate. Basically, the neural networks learn to recognize patterns and assess the correct probability of success.

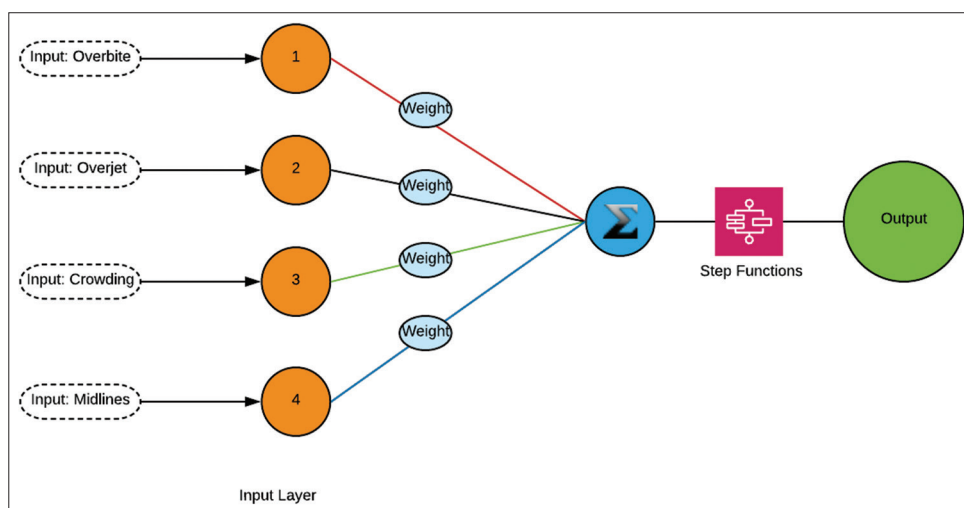
The simplest neural network consists of two layers created by a vertical arrangement of nodes and connectors forming one input and one output.<sup>[32]</sup> The nodes can perform simple data processing through arithmetic functions.<sup>[33]</sup> The connectivity between the nodes gives the power to the network. The color of the connectors represents the “weight” or the influence of the given node in the network [Figure 2].

The network then establishes the output or the most probable diagnosis. This output is dependent on the amount of data (input) and the weights allocated to the data at the input level. For neural networks to learn and be efficient, very large amount of data must be correctly labeled and weighted [Figure 3].

Finding and applying the most appropriate treatment approach to a specific malocclusion despite the multiple variables presented requires that the software accumulates and trains on a large quantity of data.

### SUPERVISED VERSUS UNSUPERVISED LEARNING

The use of AI in Orthodontics is, for the moment, limited to supervised learning such as objects or point recognition. The best example is cephalometric software programs such as WEBceph™ or AudaxCeph™, Cephio™, CephX™, DentalIQ-Ortho™, EYES.OF.AI™, and FPT-Software™ which are trained to recognized points in cephalometric radiographs to facilitate cephalometric analyses.<sup>[34]</sup> For clinical purposes, Cephalometric Landmark Identification Data could readily be



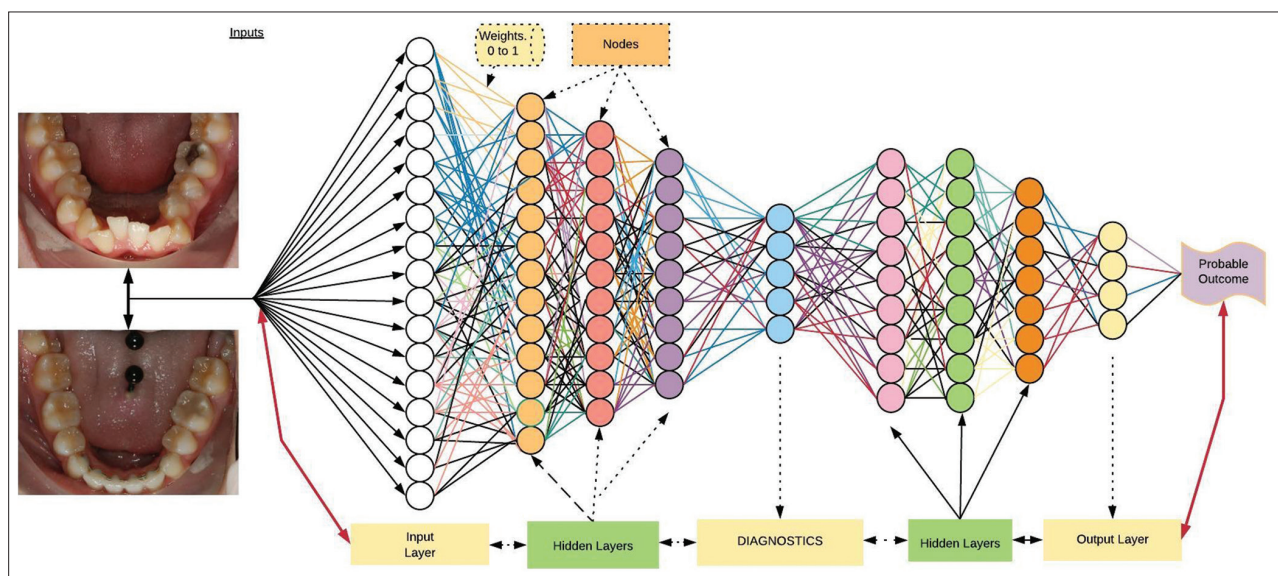
**Figure 2:** Simple neural network showing the input layer, the hidden layer, and the output layer. The lines between the input layer and the hidden layer are of different colors illustrating different weights applied.

extended even to predict and visualize soft-tissue changes after treatment. The application of AI in automated cephalometric landmark identification may lessen the burden and alleviate human errors. By gathering radiographic data automatically, it also helps reduce human tasks and the time required for both research and clinical purposes (Ji-Hoon Park *et al.*). AI showed as accurate an identification of cephalometric landmarks as did human examiners. AI always detects identical position which implies that AI may be the reliable option for repeatedly identifying multiple cephalometric landmarks [Figure 4] (Hye-Won Hwang *et al.*).

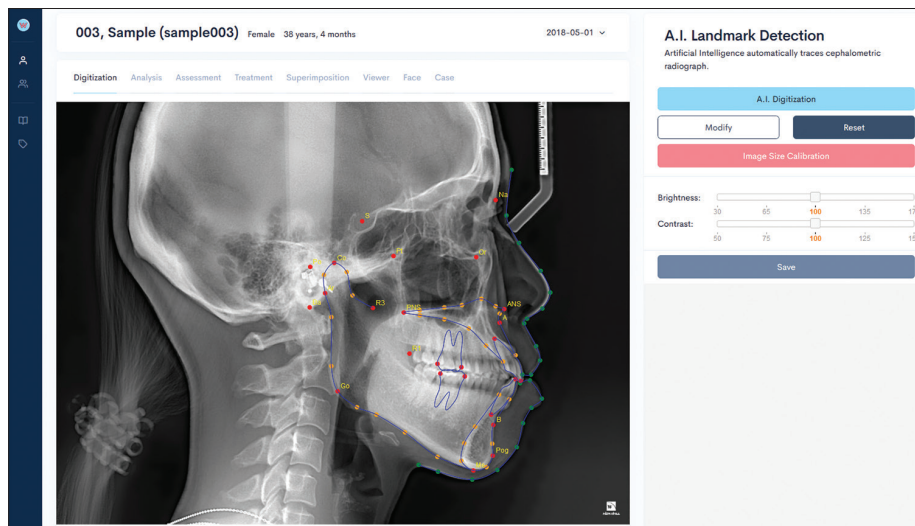
Another use of AI is found in the automation of case setups for indirect bracketing or production of aligners. These

automated setups allow for the visualization of treatment outcomes using specific algorithms again based on supervised learning. More efficient biomechanical applications using 3D physics are already reported.<sup>[35,36]</sup> New algorithms are being developed to identify teeth in Cone Beam CTs automatically [Figure 5].<sup>[37]</sup>

Unsupervised learning is referred to a higher level of complexity where the data provided is not labeled or classified. The computer program in an unsupervised mode will “train” algorithms to recognize patterns and suggest potential outcomes from a dataset.<sup>[38]</sup> In orthodontics, unsupervised learning is starting to be used to input a large quantity of malocclusions and let the computer predict the



**Figure 3:** Deep neural network. The variables on the left are input into nodes. Random weights are allocated to the layers and allow the network to “learn” by back propagation.



**Figure 4:** Cephalometric automatic point detection.

most appropriate treatment options. The before and after treatment casts of a large number of orthodontic cases are fed into a neural network. The neural network will then recognize patterns and train itself to recognize and suggest the best course of action.<sup>[39]</sup> Another application is to assist the orthodontist in choosing a modality of treatment such

as an extraction or a non-extraction pattern.<sup>[40]</sup> All reports are predictive in essence totally, but the authors warn that the validity of their findings is dependent on the training received by the neural network. This network would lose predictive accuracy rapidly if the case studied is not readily identified in the training dataset [Figure 6a and b].<sup>[39]</sup>

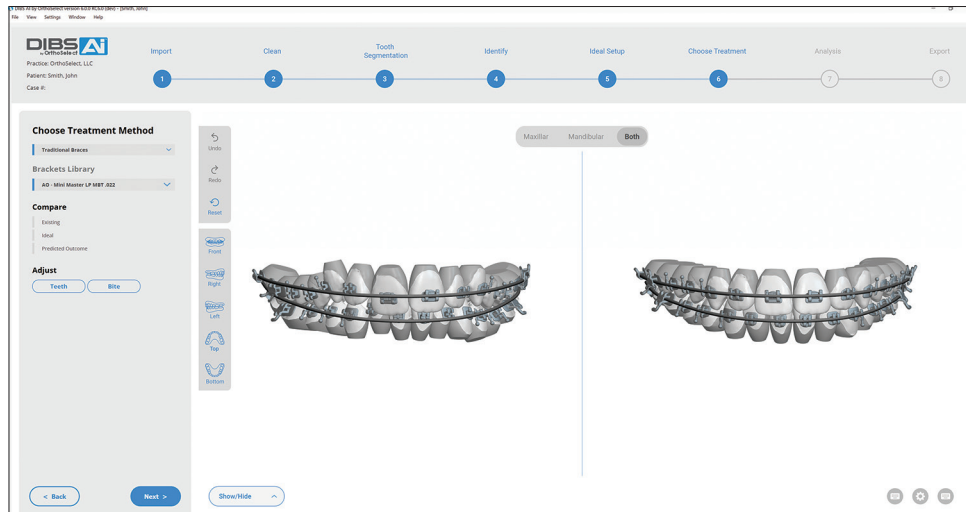


Figure 5: Indirect bracketing using artificial intelligence.

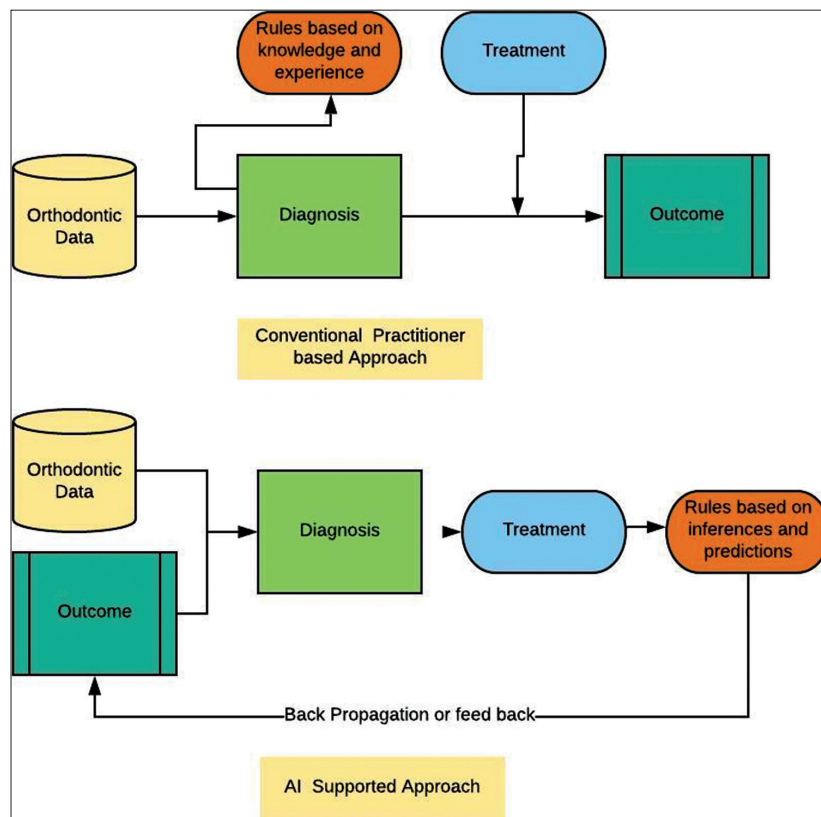


Figure 6: (a) Conventional approach to diagnose and treatment plan an orthodontic case. (b) Artificial intelligence supported approach to diagnose and treatment plan orthodontic cases.

## What can our profession do to prepare and be part of the upcoming AI revolution?

“Data is power” in the digital age and it becomes evident that solutions must be put forward to ensure that orthodontic data are protected by the profession in the same manner as medical data are not disseminated without previous understanding of its potential use.<sup>[41]</sup> When shared and protected adequately, the amount of data gathered each day in orthodontic offices combined to deep neural networks have the potential to become a great source of advancement for the profession. These neural networks will first need to be proven to be functional and reliable.

Big data feed deep learning machines. To be efficient and of use to the orthodontic field, significant amount of data needs to be gathered and processed. Malocclusions, with their multitude of variables presented to the orthodontist, require very complex networks before machine learning using unsupervised learning will have a significant impact on the treatment of malocclusions.

Even if some of the reports on the accomplishments of the latest algorithms seem revolutionary, the applications for this technology are narrow and difficult to implement. It will be up to the orthodontic profession to adapt to this new and highly disturbing environment. Currently, the bulk of the research is performed by companies leaving the orthodontic profession in a potential vulnerable position.

### Declaration of patient consent

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

### Financial support and sponsorship

Nil.

### Conflicts of interest

There are no conflicts of interest.

## REFERENCES

- Joffe L. Invisalign: Early experiences. *J Orthod* 2003;30:348-52.
- Melkos AB. Advances in digital technology and orthodontics: A reference to the Invisalign method. *Med Sci Monit* 2005;11:PI39-42.
- Charalampakis O, Iliadi A, Ueno H, Oliver DR, Kim KB. Accuracy of clear aligners: A retrospective study of patients who needed refinement. *Am J Orthod Dentofacial Orthop* 2018;154:47-54.
- Madnick SE. Information Technology Platform for the 1990's; 1989.
- Morton J, Matov V, Cao H, Kimura R, Cheng J, Pesenti B. Orthodontic Systems and Methods Including Parametric Attachments, Google Patents; 2015.
- Moorrees CF, Grøn AM. Principles of orthodontic diagnosis. *Angle Orthod* 1966;36:258-62.
- Ackerman JL. Orthodontics: Art, science, or trans-science? *Angle Orthod* 1974;44:243-50.
- Tarvit DJ, Freer TJ. Assessing malocclusion-the time factor. *Br J Orthod* 1998;25:31-4.
- Proffit WR. The evolution of orthodontics to a data-based specialty. *Am J Orthod Dentofacial Orthop* 2000;117:545-7.
- Brightman BB, Hans MG, Wolf GR, Bernard H. Recognition of malocclusion: An education outcomes assessment. *Am J Orthod Dentofacial Orthop* 1999;116:444-51.
- Gelfand AE, Hills SE, Racine-Poon A, Smith AF. Illustration of bayesian inference in normal data models using gibbs sampling. *J Am Stat Assoc* 1990;85:972-85.
- Rohrer B. End-to-end Machine Learning Library; 2020. Available from: <https://www.e2eml.school/blog.html>.
- Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med* 2003;78:775-80.
- Thanathornwong B. Bayesian-based decision support system for assessing the needs for orthodontic treatment. *Health Inform Res* 2018;24:22-8.
- 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.
- Norman GR, Eva KW. Diagnostic error and clinical reasoning. *Med Educ* 2010;44:94-100.
- Einhorn HJ, Hogarth RM. Prediction, diagnosis, and causal thinking in forecasting. In: *Behavioral Decision Making*. Berlin: Springer; 1985. p. 311-28.
- Merrifield LL, Klontz HA, Vaden JL. Differential diagnostic analysis system. *Am J Orthod Dentofacial Orthop* 1994;106:641-8.
- Sultan S, Hynes N. Multilayer flow modulator stent technology: A treatment revolution for US patients? *Expert Rev Med Devices* 2015;12:217-21.
- McNamara JA Jr. Ordinary orthodontics: Starting with the end in mind. *World J Orthod* 2000;1:45-54.
- Kohavi R, John G, Long R, Manley D, Pfleger K, editors. *MLC++: A machine learning library in C++*. In: *Proceedings Sixth International Conference on Tools with Artificial Intelligence TAI 94*. United States: IEEE; 1994.
- Chabert JL, Barbin É. *A History of Algorithms: From the Pebble to the Microchip*. Berlin: Springer; 1999.
- Mayer-Schönberger V, Cukier K. *Big Data: A Revolution that Will Transform How We Live, Work, and Think*. Boston, MA: Houghton Mifflin Harcourt; 2013. p. 242.
- Ossi Y, Jari P. Perspectives to definition of big data: A mapping study and discussion. *J Innov Manag* 2016;4:69-91.
- LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436-44.
- Grischke J, Johannsmeier L, Eich L, Griga L, Haddadin S. Dentronics: Towards robotics and artificial intelligence in dentistry. *Dent Mater* 2020;36:765-78.
- Krizhevsky A, Sutskever I, Hinton GE. Image nzet classification with deep convolutional neural networks. *Commun ACM* 2017;60:84-90.

28. Ghahramani Z. Probabilistic machine learning and artificial intelligence. *Nature* 2015;521:452-9.
29. Koller D, Friedman N. Probabilistic Graphical Models: Principles and Techniques. Cambridge: MIT Press; 2009.
30. Jordan MI. An Introduction to Probabilistic Graphical Models, preparation; 2003.
31. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, *et al.* Backpropagation applied to handwritten zip code recognition. *Neural Comput* 1989;1:541-51.
32. Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain. In: *Neurocomputing: Foundations of Research*. Cambridge: MIT Press; 1988. p. 89-114.
33. Meyer-Baese A. *Pattern Recognition in Medical Imaging*. United States: Academic Press, Inc.; 2003.
34. Rao GK, Mokhtar N, Iskandar YH, Srinivasa AC, editors. Learning orthodontic cephalometry through augmented reality: A conceptual machine learning validation approach. In: *2018 International Conference on Electrical Engineering and Informatics (ICELTICs)*. United States: IEEE; 2018.
35. Morton J, Derakhshan M, Kaza S, Li C, editors. Design of the invisalign system performance. In: *Seminars in Orthodontics*. Amsterdam: Elsevier; 2017.
36. Stapleton BM, Lin WS, Ntounis A, Harris BT, Morton D. Application of digital diagnostic impression, virtual planning, and computer-guided implant surgery for a CAD/CAM-fabricated, implant-supported fixed dental prosthesis: A clinical report. *J Prosthet Dent* 2014;112:402-8.
37. Cui Z, Li C, Wang W, editors. Tooth net: Automatic tooth instance segmentation and identification from cone beam CT images. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2019.
38. Zhou C, Pan S, Zhou T, editors. Design and implementation of software simulation system for dental orthodontic robot. In: *IOP Conference Series: Materials Science and Engineering*. Bristol: IOP Publishing; 2019.
39. Wei G, Cui Z, Liu Y, Chen N, Chen R, Li G, *et al.* TANet: Towards Fully Automatic Tooth Arrangement. Cham: Springer; 2020.
40. 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.
41. Mostert M, Bredenoord AL, Biesart MC, van Delden JJ. Big data in medical research and EU data protection law: Challenges to the consent or anonymise approach. *Eur J Hum Genet* 2016;24:956-60.

**How to cite this article:** Retrouvey JM. The role of AI and machine learning in contemporary orthodontics. *APOS Trends Orthod* 2021;11(1):74-80.