

Original Article

## Automated artificial intelligence modeling for clinical decision-making of extraction versus non-extraction cases: Unleashing the potential for all

Priyanka Jain<sup>1</sup>, Gururaj Suresh Patil<sup>2</sup>, Sayeeda Laeque Bangi<sup>1</sup>, Prasad Konda<sup>1</sup>, Safiya Sana<sup>3</sup>, Revati Patil<sup>1</sup>

<sup>1</sup>Department of Orthodontics and Dentofacial Orthopaedics, Al-Badar Rural Dental College and Hospital, Gulbarga, Karnataka, India, <sup>2</sup>Department of Computer Science, Santa Clara University, Santa Clara, California, United States, <sup>3</sup>Department of Orthodontics, ESIC Dental College, Gulbarga, Karnataka, India.



**\*Corresponding author:**

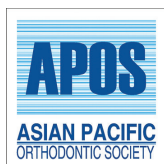
Priyanka Jain,  
Department of Orthodontics  
and Dentofacial Orthopaedics,  
Al-Badar Rural Dental College  
and Hospital, Gulbarga,  
Karnataka, India.

[pinksssjain10@gmail.com](mailto:pinksssjain10@gmail.com)

Received: 29 May 2025  
Accepted: 27 November 2025  
Epub Ahead of Print: 23 March 2026  
Published:

DOI  
[10.25259/APOS\\_130\\_2025](https://doi.org/10.25259/APOS_130_2025)

Quick Response Code:



### ABSTRACT

**Objectives:** To develop and evaluate artificial intelligence-based machine learning models for informed decision-making between extraction and non-extraction orthodontic treatment scenarios, recognizing the irreversible nature of extractions and the potential for adverse treatment outcomes if clinical decisions are made inaccurately.

**Material and Methods:** Eight hundred patients (18–35 years old) with lateral cephalograms and study models were gathered from the department record room. The dataset was split into a training set and a test set to train the model by machine programming to evaluate the accuracy, recall, and precision performances of three distinct models toward decision-making of extraction versus non-extraction cases, respectively.

**Results:** The success rates of the classifier system for the decision tree classifier were 87%, the random forest classifier was 97%, and the XGB classifier was 98%. The XGB classifier had the highest accuracy, recall, and precision scores, exhibiting outstanding reliability.

**Conclusion:** The success rates of the classifier algorithms for the diagnosis of extraction versus non-extraction suggested that AI expert systems with ML algorithms could be a new approach in orthodontic treatment planning and clinical decision-making.

**Keywords:** Artificial intelligence, Extraction, Machine learning, Non-extraction

### INTRODUCTION

With continued global research efforts, technology has advanced significantly. The past 25 years have seen a fourfold expansion in the field of information technology, and its clinical use in dentistry has had a considerable impact on both general and orthodontic dentistry.<sup>[1]</sup> Artificial intelligence (AI), machine learning (ML), robotics, and other technologies have exceedingly decreased the amount of time needed, required human prowess, etc. Researchers and physicians in all domains can benefit from its sophisticated pattern identification and prediction algorithms.<sup>[2-4]</sup> Originally, attempts were made to create handcrafted AI, which is the process of trying to fit every possible scenario and its related solution into a computer program. In some situations, it functions well. However, because AI is incapable of “learning” anything on its own, it is by no means capable of surpassing the human brain. Regarding contemporary AI, which aims to give machines the capacity to learn, it has the capacity to acquire capabilities

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, transform, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

©2026 Published by Scientific Scholar on behalf of APOS Trends in Orthodontics

that surpass those of its developer. ML is how we attempt to achieve AI.<sup>[5]</sup>

In general, ML is the process of deriving a function from input data. Later, the function can convert fresh incoming data – such as sounds and images – into useful output data, such as facial and speech recognition. One may categorize ML into two processes: Testing and training. During the training process, a collection of functions, referred to as a model, is assessed using computer algorithms and training data to determine the optimal function. After the testing procedure, the selected function is evaluated to decide whether or not it is feasible by reevaluating the testing data.<sup>[5]</sup>

Many areas of orthodontics can benefit from the application of AI (artificial learning), ML, and deep learning, given the quickly expanding and developing disciplines of technological advancements.<sup>[6]</sup> When used in orthodontics, AI and ML systems offer useful instruments that can enhance clinical procedures. Orthodontists practicing with these clinical decision support tools can work more productively, with less subjectivity and variability.<sup>[7,8]</sup> Digital data processing, automated cephalometric tracing, diagnosis, and treatment planning, predicting the need for orthodontic extractions, facial image recognition, a framework for automated diagnosis and computer-assisted planning in plastic and reconstructive surgeries, identifying the cervical vertebrae maturity stages for growth prediction, and more are among the applications of AI in the orthodontic specialty.<sup>[9]</sup> Rather than focusing solely on the oral and maxillofacial region, orthodontists must make a holistic evaluation of their patients by examining various facial structures from various angles.

To increase evaluation speed and accuracy, diagnostic imaging must be automated.<sup>[11]</sup> A proper diagnosis and carefully thought-out treatment procedures are essential for the effectiveness of orthodontic therapy. Making a treatment plan is the most crucial step in orthodontic therapy. Choosing which teeth to take and whether to perform an extraction is a crucial step in treatment planning because extractions cannot be reversed.<sup>[11]</sup> Based on their expertise and experience, orthodontists typically use information from radiography, dental models, photos, and clinical evaluations to determine the best course of action. In the process of treatment planning, this frequently results in intra-clinician and inter-clinician variability. Treatment plans may differ depending on which data were utilized for the diagnosis. Decisions on extraction that differ could be significant.<sup>[10]</sup>

Automated systems based on AI can be very helpful in clinical decision-making and serve as a backup for physicians, particularly those with limited clinical expertise.<sup>[11]</sup> This study was conducted to prepare AI decision-making models of extraction versus non-extraction cases and to compare the performance of three different ML classifier

algorithms, that is, random forest classifier, decision tree classifier, and Extreme gradient boost (XGBoost) classifier, to provide researchers and clinicians with powerful pattern-finding algorithms.

## MATERIAL AND METHODS

### Study design

About 800 sample records were recruited from patients who visited or visiting the department of orthodontics with complaints such as irregularly placed teeth, dental arch constrictions, incompetent lips, and unpleasant facial appearances. This study has been approved by the Institutional Ethical Committee. Records such as lateral cephalograms and study models were collected from the departmental record room.

### Inclusion criteria and exclusion criteria

The inclusion criteria were patients aged between 13 and 35 years having crowding >5 mm with proclined upper and lower anteriors, full permanent dentition, increased overjet, and overbite. The exclusion criteria were patients having congenital anomalies, syndromes, partial anodontia, growing patients, and the presence of pathologic diseases.

### Methodology

From 800 samples, 600 subjects were assigned to the learning set, and 200 subjects were assigned to the test set. Cases were categorized into extraction and non-extraction groups based on model analysis and incisor proclination. Borderline cases were assigned to the non-extraction group if significant interdental spacing was observed or if arch expansion was deemed feasible, whereas incisor proclined cases with minimal interdental space availability and well-aligned basal arches were allocated to the extraction group. Robust data preprocessing steps were applied, including noise filtering, normalization, and accurate label assignment, to enhance model performance and minimize data-related biases, such that only those cases were selected that had a good outcome to clean the model or label the data having an unacceptable outcome. The test set was used only for the evaluation of the models. Next, through evaluation of the test set, adequacy and accuracy were evaluated, and the best-fit model was chosen. The models were prepared using three supervised ML classifier algorithms, that is, decision tree making, random forest classifier, and XGBoost classifier for decision-making of extraction versus non-extraction cases. Further, the comparison of the performance of three different ML algorithms was done to evaluate the accuracy of the three different models. Each model was trained independently to generate the result.

Nineteen measurements (angular and linear hard and soft-tissue lateral cephalogram parameters and study model parameters) [Tables 1-3] were selected for the input data: SNA, SNB, ANB, Witt’s appraisal, FMA, IMPA, β angle, Y-axis, angular UI-NA, linear UI-NA, angular LI-NB, linear LI-NB, nasolabial angle, nasio-mental angle, E line-upper lip, E line-lower lip, Bolton’s ratio, Ashley’s model analysis, and Nance model analysis. These parameters had strong clinical significance and relevance in assessing cases requiring extractions or not, from anteroposterior relationships, vertical relationships, tooth inclinations, and soft-tissue characteristics. Extraction and non-extraction diagnosis of the total data was performed by the construction of the classifier model. In comparison with an actual diagnosis given by different orthodontists, the decision-making success rate was calculated. Finally, the total success rate of the diagnosis of extractions and non-extractions was calculated for three different models, and the comparison of their performance was done.

**RESULTS**

To understand the true utility of these models in a clinical setting, the ability of each model to correctly predict the treatment modality, referred to as model “accuracy,” must be determined. The confusion matrix [Figure 1] was used to calculate the accuracy of the performance of each model in diagnosing extraction and non-extraction cases for all data sets. Accuracy was computed from the confusion matrix for each test set, and the average accuracy across all test sets was reported for each model.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

- TP: True Positives
- TN: True Negatives
- FP: False Positives
- FN: False Negatives

Accuracy is calculated as the ratio of true positives and true negatives to the total number of predictions (true positives, true negatives, false positives, and false negatives)<sup>[12]</sup>

The obtained confusion matrices for extraction versus non-extraction cases by implementing the decision tree classifier [Figure 1a], random forest [Figure 1b], and XGBoost classifier algorithms [Figure 1c] for the overall dataset suggested that the decision tree classifier correctly predicted 76 non-extraction and 64 extraction instances but misclassified nine non-extraction cases as extraction and 10 extraction cases as non-extraction. This reflected a moderate performance with notable room for improvement in both sensitivity and specificity. The random forest classifier showed a marked improvement, correctly identifying 82 non-extraction and

**Table 1:** Cephalometric parameters for hard tissues.

Measurements	Description	Construction
SNA	Angle between sella-nasion and nasion-point A lines	Sagittal
SNB	Angle between sella-nasion and nasion-point B lines	Sagittal
ANB	Angle between nasion-point A and nasion-point B lines	Sagittal
Wits appraisal	Perpendiculars from points A and B to occlusal plane (points AO and BO)	Sagittal
β angle	Angle between A-B line and a perpendicular from point A to condylar axis	Sagittal
FMA	Angle between Frankfort horizontal and Tweed’s mandibular plane	Vertical
Y-axis	Angle between Frankfort horizontal plane and sella-gnathion line	Vertical
UI-NA angular	Angle between long axis of upper incisor and nasion-point A line	Angular
UI-NA linear	Distance from upper incisor tip to nasion-point A line	Linear
LI-NB angular	Angle between long axis of lower incisor and nasion-point B line	Angular
LI-NB linear	Distance from lower incisor tip to nasion-point B line	Linear
IMPA	Angle between central axis of lower incisor and Tweed’s mandibular plane	Angular

73 extraction cases with only four misclassifications in total (three false positives and one false negative). This indicated enhanced precision and recall, especially in detecting extraction cases more accurately. The XGBoost classifier demonstrated the best classification accuracy among the three. It correctly predicted all 74 extraction cases and 82 out of 85 non-extraction cases, with just three false positives and zero false negatives. This perfect recall for the extraction class, combined with minimal error, highlighted XGBoost’s superior learning and generalization capabilities, offering significantly higher reliability and classification precision, making it more suitable for critical decision-making applications.

The bar plot [Figure 2] illustrated the performance metrics of each model. The XGB classifier had the highest accuracy score of 0.985, followed by the random forest classifier and decision tree classifier with 0.975 and 0.881, respectively.

The XGB classifier and random forest classifier had a similar precision score of 0.961, and the decision tree classifier had a 0.877 precision score. The XGB classifier had the highest recall score of 1, followed by the random forest and decision tree classifiers with 0.986 and 0.865, respectively. The F1 score was highest for the XGB classifier, followed by the random forest and decision tree classifier, being 0.98, 0.973, and 0.871, respectively. The above findings suggested that the XGB classifier algorithm outperformed the random forest classifier and decision tree classifier, achieving the highest recall score (1.00) and overall balanced performance.

[Figure 3] displayed the Pearson correlation coefficients among various cephalometric measurements, revealing the strength and direction of linear relationships. The matrix aided in identifying interdependent features that may influence diagnostic or treatment decisions in orthodontic analysis. Each cell in the heatmap showed a Pearson correlation coefficient ranging from -1 to +1, indicating the direction and strength of correlation between two variables. A value closer to +1 indicates a strong positive linear relationship, while

values near 0 indicate weak or no correlation. SNA and SNB showed a strong positive correlation (0.69), as both measure cranial base angles relative to the maxilla and mandible, respectively. ANB and Witts appraisal are highly correlated (0.63), reflecting their shared role in assessing anteroposterior jaw discrepancy. UI-NA (deg) and UI-NA (mm) showed a strong correlation (0.5), indicating that changes in angular measurements often accompany proportional changes in linear distance. LI-NB (deg and mm) also exhibited a high correlation (0.51), consistent with observations in lower incisor analysis. Ashley’s and Nance’s analyses, which assess dental arch relationships, were significantly correlated (0.52), and both were highly correlated with Bolton’s ratio (Ashley’s: 0.53; Nance: 0.66), suggesting their mutual importance in evaluating dental proportionality. IMPA, which reflects incisor mandibular plane angle, showed moderate correlations with both UI-NA (deg: 0.28) and LI-NB (deg: 0.48), suggesting its influence on incisor inclination.

The performance error was also calculated for all three algorithm types between training and testing. The learning curves for the decision tree, random forest, and XGBoost classifiers revealed distinct trends in how each model generalizes with increasing training data. The decision tree classifier [Figure 4a] showed consistently zero error on the training set, indicative of overfitting. The test set error remained relatively high and fluctuated with different training set sizes, suggesting poor generalization and high variance. While there was a slight downward trend in test error, the lack of convergence indicates limitations in model complexity and stability. In contrast, the random forest classifier [Figure 4b] demonstrated better generalization. The training error was zero, similar to the decision tree, but the test error was significantly lower and stabilized early as more training data were added. The minimal variance in test error across all training sizes reflects the ensemble model’s robustness and ability to prevent overfitting by aggregating multiple decision trees. The XGBoost classifier [Figure 4c]

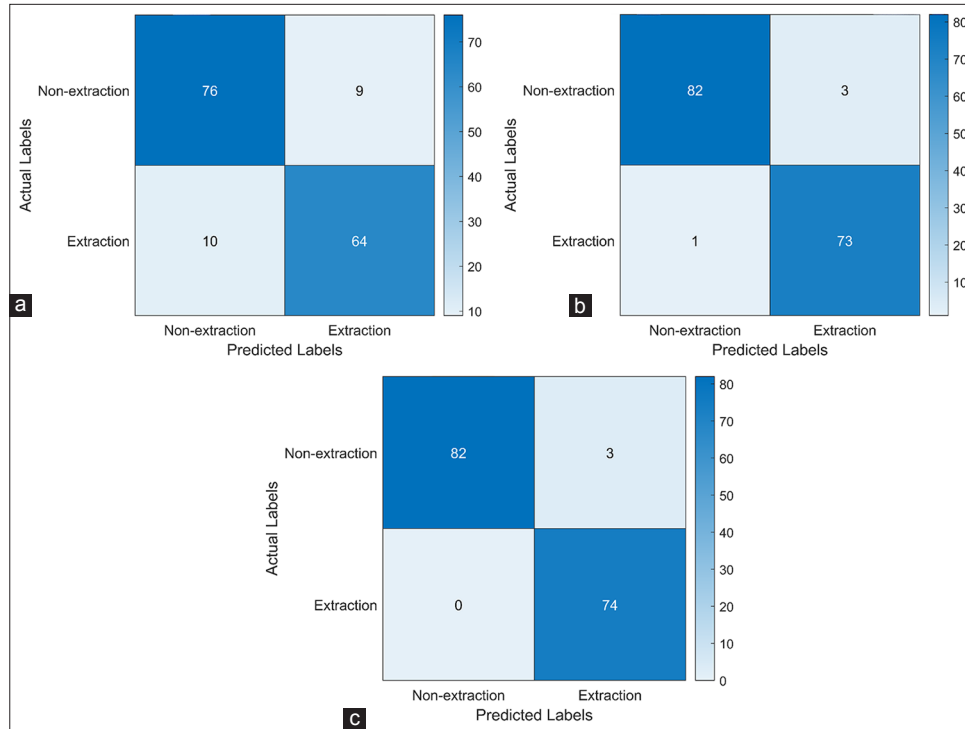
**Table 2:** cephalometric parameters for soft tissues.

Measurements	Description	Construction
Nasiomental angle	Angle between nasion–nasal dorsum line and nasal tip–soft tissue pogonion line	Sagittal
Nasolabial angle	Angle between PCm tangent and PCm–Ls line; formed at the junction of nose base and upper lip	Sagittal
E line-upper lip	Distance from upper lip to the esthetic line (pronasale to soft tissue pogonion)	Sagittal
E line-lower lip	Distance from lower lip to the esthetic (Ricketts) line	Sagittal

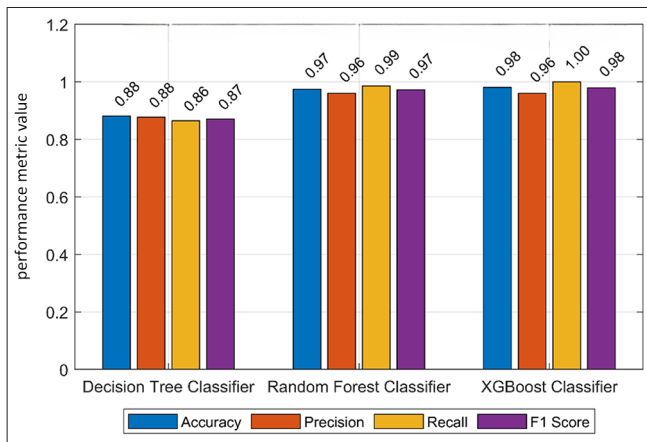
**Table 3:** parameters for study model analysis.

Analysis	Description	Interpretation
Bolton’s analysis	Assesses tooth size discrepancy between mandibular and maxillary teeth. Ideal ratios: 77.2% (anterior), 91.3% (overall).	Ratio <91.3%: Maxillary excess (overall) Ratio >91.3%: Mandibular excess (overall) Ratio <77.2%: Maxillary anterior excess Ratio >77.2%: Mandibular anterior excess
Ashley Howe’s analysis	Evaluates relationship between tooth material (16–26) and basal arch width (14–24); used to assess feasibility of arch expansion.	Ratio <37%: Extraction case Ratio 37–44%: Borderline case Ratio >44%: Non-extraction case P.M.D >P.M.B.A.W: No expansion P.M.D <P.M.B.A.W: Expansion possible
Nance and Carey’s index	Determines arch length discrepancy using: LD=LA+2X, where LA=sum of incisor widths, X=sum of widths of teeth 3,4,5.	<2.5mm: non-extraction case 2.5mm-5mm: second premolar extraction>5mm: first premolar extraction

P.M.D: Premolar diameter, P.M.B.A.W: Premolar basal arch width, LD: Linear dimension, LA: Sum of incisor widths



**Figure 1:** (a) Confusion matrix for a decision tree classifier showing classification performance. (b) Confusion matrix for a random forest classifier demonstrating classification accuracy. (c) Confusion matrix for an extreme gradient boost classifier showing optimal classification performance.



**Figure 2:** Bar plot of performance metrics of the three models.

further improved on this trend, showing the lowest and most stable test error across the entire range of training set sizes. Like the others, it also achieved zero training error, but unlike the decision tree, its test error consistently decreases and levels off with more data. This indicated that XGBoost achieved both high accuracy and excellent generalization due to its use of boosting, regularization, and optimization techniques. Overall, while the decision tree struggles with overfitting and generalization, Random Forest offers more stable performance, and XGBoost clearly outperforms both

in terms of predictive consistency and learning efficiency.

The receiver operating characteristic (ROC) curves illustrated the classification performance of three different ML models: Each curve plots the true-positive rate (sensitivity) against the false-positive rate, and the area under the curve (AUC) value that quantifies overall model performance. The decision tree classifier [Figure 5a] showed a moderately strong performance with an AUC of 0.88, indicating its ability to distinguish between classes, though with a relatively higher false-positive rate. In contrast, the random forest classifier [Figure 5b] achieved a significantly higher AUC of 0.97, reflecting improved sensitivity and specificity due to its ensemble-based architecture. The XGBoost classifier [Figure 5c] outperformed both, with an AUC of 0.98, closely approaching an ideal classifier. This highlights its superior learning capacity, optimization, and regularization techniques, which contribute to minimal misclassifications. Overall, while the decision tree provides a baseline level of accuracy, the random forest and especially the XGBoost classifiers deliver enhanced and more reliable performance in terms of true positive recognition and reduced false alarms.

## DISCUSSION

ML has been applied to categorization and decision-making difficulties in numerous research.<sup>[13-16]</sup> One way to approach

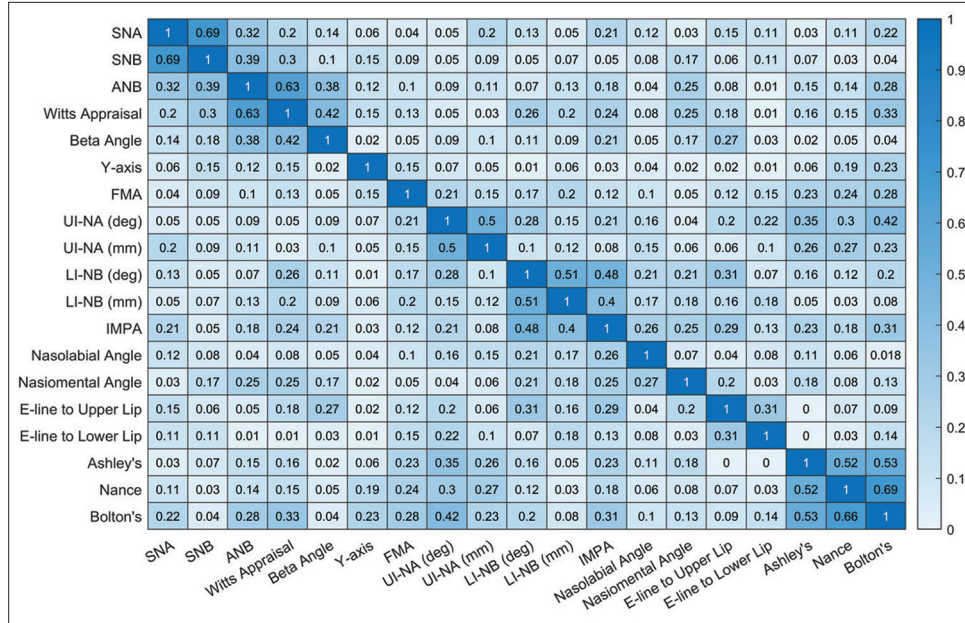


Figure 3: Heatmap of Pearson correlation matrix for cephalometric and dental variables.

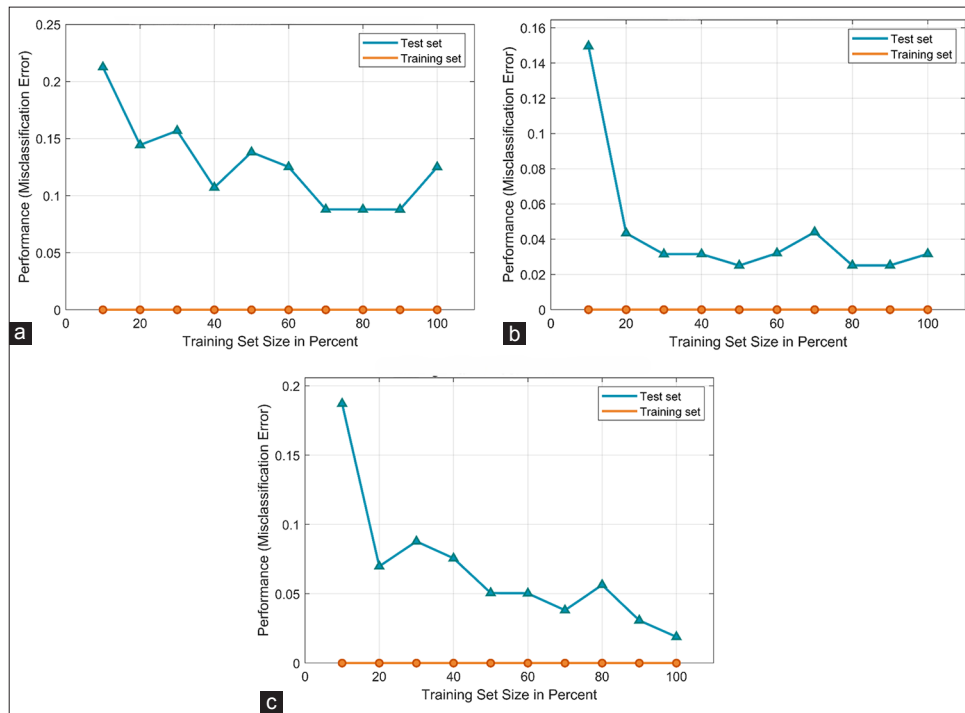
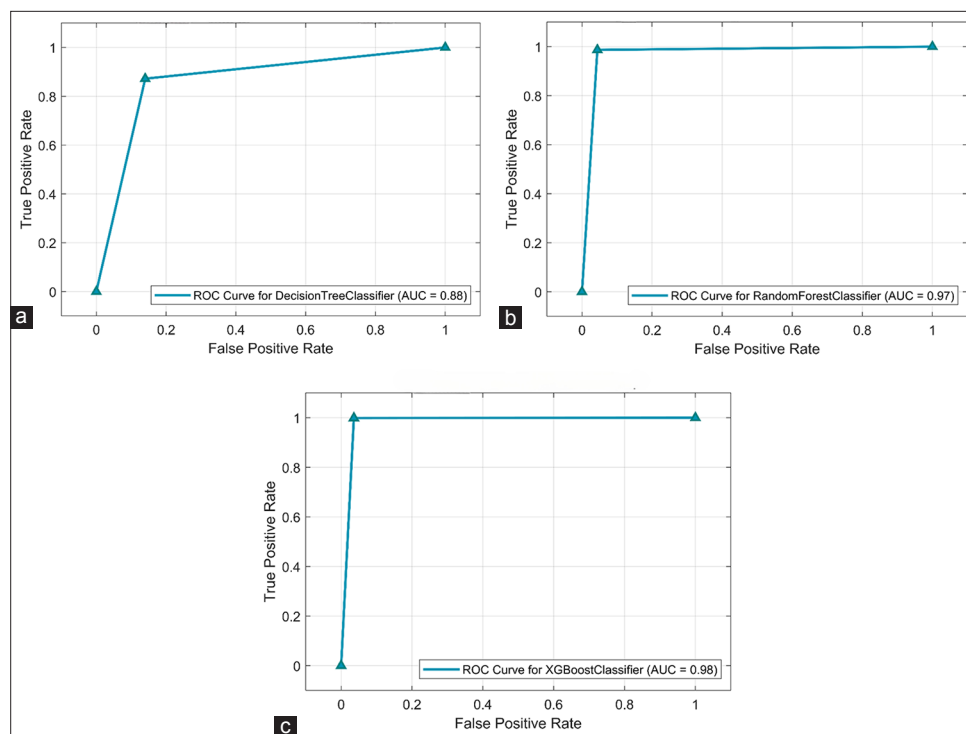


Figure 4: (a) Learning curve showing misclassification error for the decision tree classifier across varying training set sizes. (b) Learning curve showing misclassification error for the random forest classifier across varying training set sizes. (c) Learning curve showing misclassification error for the extreme gradient boost classifier across increasing training set sizes.

the decision to extract is a sort of categorization problem. Over the past several years, the number of scoping reviews about “AI and ML algorithms and their application in orthodontics” that have been published in orthodontic literature has

steadily increased.<sup>[17,18]</sup> Numerous studies demonstrate how well various types of algorithms may be applied to segment, automatically detect, analyze, and extract characteristics from images to improve orthodontic diagnosis.<sup>[10,19-25]</sup>



**Figure 5:** (a) The receiver operating characteristic (ROC) curves illustrating the classification performance of the decision tree classifier. (b) The ROC curves illustrate the classification performance of the random forest classifier. (c) The ROC curves illustrate the classification performance of the extreme gradient boost classifier. AUC: Area Under Curve.

AI models are increasingly transforming clinical decision-making by providing data-driven support that enhances diagnostic accuracy, optimizes treatment planning, and improves patient outcomes. Their integration into clinical workflows allows for real-time interpretation of complex datasets, leading to more precise and customized treatment plans and acceptable outcomes. AI also plays a pivotal role in risk stratification and prognosis prediction in orthognathic jaw surgery cases. The direct clinical impact of AI lies in its ability to augment human expertise, reduce diagnostic and therapeutic variability, and facilitate more informed, consistent, and efficient decision-making. As these technologies continue to evolve, their responsible implementation holds the promise of advancing precision and elevating standards of care.

In general, extractions are done to relieve crowding, lessen the prominence of dental arches, and adjust the disparity in the anteroposterior connections between arches. In addition, spaces created by tooth extraction enable the correction of vertical dimension disparities. Moreover, extraction spaces can be used to rectify tooth size discrepancies (congenital agomphosis or Bolton index abnormality), width discrepancies across arches (cross-bite or scissor bite of posterior teeth), and so on.<sup>[19]</sup> Malocclusions were considered to have “improved” if the PAR index values were lower after therapy

than they were before treatment by more than 30%.<sup>[26]</sup> Choosing between extraction and non-extraction therapy can be one of the most difficult daily decisions in a clinical orthodontic practice, particularly in situations that are considered “orderline.”

To create a highly precise simulation performance model and ensure a successful course of therapy, the present study employed more significant parameters. The study utilizes widely recognized ML algorithms; however, its innovation resides in the clinical design architecture that integrates the largest multivariable cephalometric-study model dataset ( $n = 800$ ) with clinically validated parameters,<sup>[19]</sup> stringent noise filtering, the inclusion of solely clinically successful outcomes, and an extensive performance comparison and generalization analysis (ROC, AUC curve, and confusion matrix). All of these things combined make a model for making irreversible extraction decisions that are more clinically dependable and easier to understand. This suggests that the developed population-specific cohort-based model, which makes use of three distinct ML algorithms, can be a useful, evidence-based tool for determining which treatment option is preferable when it comes to extractions versus arch expansion.<sup>[10,27,28]</sup>

In borderline orthodontic cases where the decision between extraction and non-extraction is not clear-cut, AI models

assist by analyzing a combination of clinical features such as crowding severity, incisor inclination, arch space, and facial profile.<sup>[28]</sup> These models are trained on large datasets that include past treatment decisions and outcomes, allowing them to recognize patterns and predict which treatment path is more likely to succeed. Rather than giving a simple yes-or-no answer, the AI often provides a probability or confidence level for each option, helping the clinician weigh the risks and benefits. In some systems, the AI also highlights which factors most influenced the recommendation, offering transparency and supporting the clinician's final judgment. This helps reduce uncertainty and supports more consistent and informed decision-making in challenging cases.

The strength of this model lies in the fact that it was created from a pool of 800 patients treated by different attending orthodontists using a variety of treatment modalities and has varying treatment philosophies, reducing the risk of institutional bias. Only cases with post-treatment outcomes classified as objectively successful were used for model labeling. The prepared model was successful in judging whether an orthodontic treatment for a particular case needs extraction or not, since this study had considered the parameters in both anteroposterior and vertical dimensions, and both hard and soft-tissue structures along with study models. The demographic variation – such as age, sex, ethnicity, growth status, and craniofacial features – can impact treatment choices; however, the current dataset was limited to people aged 13–35 years with comparable development trends.

Although this study was successful in diagnosing cases requiring extraction or not using three different algorithm classifiers, Random Forest, XGB classifiers, and a decision tree classifier, it reduced subjectivity and variability. With high-performance metrics like an AUC of 0.98 – can greatly aid clinical decision-making, they do not account for the ethical dimensions involved in irreversible procedures such as tooth extractions. Relying too heavily on ML outputs without any critical appraisal, especially in borderline cases without fully considering the broader clinical, psychological, and long-term implications for the patient, could undermine the individualized, patient-centered approach that is foundational to ethical practice. An openly accessible, simplified inference-only version of the trained model, along with anonymized example data for independent benchmarking by researchers and clinicians, is currently under development. This open validation will ensure transparency, reproducibility, and clinical credibility.

#### Limitations and future research direction

This study did not consider the molar relationship in the sagittal dimension, as it did not extend the development of the model to determine the anchorage pattern for extraction

space closure, defining the molar relationship at the end of the treatment. Exclusion of surgical cases might limit the generalization of this model. To plan to improve the performance of the models, more development is required to diagnose surgical cases, extraction and anchorage patterns, predict treatment time, relapse risk, or the success of retention protocols. Future studies with a multi-center collaboration are required to build a cross-ethnic and demographic generalizable decision model with prediction stability.

## CONCLUSION

The following conclusions could be drawn from this research:

1. The constructed AI model in this study was able to decide the orthodontic cases requiring extraction or not with an accuracy of 98%
2. The XGB classifier system had the highest accuracy, recall, and precision scores, outperforming the random forest and decision tree classifier algorithms. However, the random forest classifier demonstrated a predictive performance comparable to the XGBoost classifier, which cannot be underestimated.
3. The use of this model could be one of the best alternatives available to simplify the process of decision-making, as it is less operator-dependent and has accurate predictions.

**Acknowledgment:** My sincere appreciation goes to my closest friend, Mr. Sharan Shastri, Ph.D., for his invaluable assistance in helping me throughout the study. Without his unwavering support, I could not have embarked on my adventure.

**Ethical approval:** The research/study was approved by the Institutional Review Board at the Institutional Ethics Committee, approval number No.IEC/2020-21/02, dated 12th February 2021.

**Declaration of patient consent:** Patient's consent was not required as there are no patients in this study.

**Financial support and sponsorship:** Nil.

**Conflicts of interest:** There are no conflicts of interest.

**Use of artificial intelligence (AI)-assisted technology for manuscript preparation:** The authors confirm that there was no use of artificial intelligence (AI)-assisted technology for assisting in the writing or editing of the manuscript, and no images were manipulated using AI.

## REFERENCES

1. Büyük SK, Hatal S. Artificial intelligence and machine learning in orthodontics. *Ortodogru Tıp Derg* 2019;11:517-23.
2. Nilsson NJ. Artificial intelligence: A new synthesis. San Francisco: Morgan Kaufmann; 1998. p. 493. Available from: <https://core.ac.uk/download/pdf/82021614.pdf> [Last accessed on 2025 May 22].
3. Faber J, Faber C, Faber P. Artificial intelligence in orthodontics. *APOS Trends Orthod* 2019;9:201-5.
4. Liu X, Sun B, Zhang Z, Wang Y, Tang H, Zhu T. Gait can reveal sleep quality with machine learning models. *PLoS One*

- 2019;14:e0223012.
5. Hung HC, Wang YC. Applications of artificial intelligence in orthodontics. *Taiwan J Orthod* 2020;32:85-9.
  6. Liu J, Zhang C, Shan Z. Application of artificial intelligence in orthodontics: Current state and future perspectives. *Healthcare (Basel)* 2023;11:2760.
  7. Khanna S. Artificial intelligence: Contemporary applications and future compass. *Int Dent J* 2010;60:269-72.
  8. Asiri SN, Tadlock LP, Schneiderman E, Buschang PH. Applications of artificial intelligence and machine learning in orthodontics. *APOS Trends Orthod* 2020;10:17-24.
  9. Sakolia A. Artificial intelligence in orthodontics. *Int J Health Sci* 2021;5:379-86.
  10. Jung SK, Kim TW. New approach for the diagnosis of extractions with neural network machine learning. *Am J Orthod Dentofacial Orthop* 2016;149:127-33.
  11. Suhail Y, Upadhyay M, Chhibber A, Kshitiz. Machine learning for the diagnosis of orthodontic extractions: A computational analysis using ensemble learning. *Bioengineering (Basel)* 2020;7:55.
  12. Kk H, Acilar AM, Izgi MS. Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. *Prog Orthod* 2019;20:41.
  13. Takada K, Yagi M, Horiguchi E. Computational formulation of orthodontic tooth-extraction decisions. Part I: To extract or not to extract. *Angle orthod* 2010;79:885-91.
  14. Yagi M, Ohno H, Takada K. Computational formulation of orthodontic tooth-extraction decisions. Part II: Which tooth should be extracted. *Angle orthod* 2009;79:892-8.
  15. Han M, Snow PB, Brandt JM, Partin AW. Evaluation of artificial neural networks for the prediction of pathologic stage in prostate carcinoma. *Cancer* 2001;91:1661-6.
  16. 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;54:122-6.
  17. Gandedkar NH, Vaid NR, Darendeliler MA, Premjani P, Ferguson DJ. The last decade in orthodontics: A scoping review of the hits, misses and the near misses! *Semin Orthod* 2019;25:339-55.
  18. Vaid NR, Hansa I, Bichu Y. Smartphone applications used in orthodontics: A scoping review of scholarly literature. *J World Fed Orthod* 2020;9:S67-73.
  19. 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.
  20. Leonardi RM, Giordano D, Maiorana F, Greco M. Accuracy of cephalometric landmarks on monitor-displayed radiographs with and without image emboss enhancement. *Eur J Orthod* 2010;32:242-7.
  21. 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;34:480-6.
  22. Patcas R, Bernini DA, 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* 2019;48:77-83.
  23. Kunz F, Stellzig-Eisenhauer A, Zeman F, Boldt J. Artificial intelligence in orthodontics: Evaluation of a fully automated cephalometric analysis using a customized convolutional neural network. *J Orofac Orthop* 2020;81:52-68.
  24. Vranckx M, Van Gerven A, Willems H, Vandemeulebroucke A, Ferreira Leite A, Politis C, *et al.* Artificial intelligence (AI)-driven molar angulation measurements to predict third molar eruption on panoramic radiographs. *Int J Environ Res Public Health* 2020;17:3716.
  25. Chen H, Zhang K, Lyu P, Li H, Zhang L, Wu J, *et al.* A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Sci Rep* 2019;9:3840.
  26. Richmond S, Shaw WC, Roberts CT, Andrews M. The PAR index (peer assessment rating): Methods to determine outcome of orthodontic treatment in terms of improvement and standards. *Eur J Orthod* 1992;14:180-7.
  27. Khanagar SB, Al-Ehaideb A, Vishwanathaiah S, Maganur PC, Patil S, Naik S, *et al.* Scope and performance of artificial intelligence technology in orthodontic diagnosis, treatment planning, and clinical decision-making - A systematic review. *J Dent Sci* 2021;16:482-92.
  28. Real AD, Real OD, Sardina S, Oyonarte R. Use of automated artificial intelligence to predict the need for orthodontic extractions. *Korean J Orthod* 2022;52:102-111.

**How to cite this article:** Jain P, Patil GS, Bangi SL, Konda P, Sana S, Patil R. Automated artificial intelligence modeling for clinical decision-making of extraction versus non-extraction cases: Unleashing the potential for all. *APOS Trends Orthod*. doi: 10.25259/APOS\_130\_2025