JAOS JOURNAL OF APPLIED ORAL SCIENCE

Original Article
http://dx.doi.org/10.1590/1678-7757-2024-0282

Using machine learning to classify temporomandibular disorders: a proof of concept*

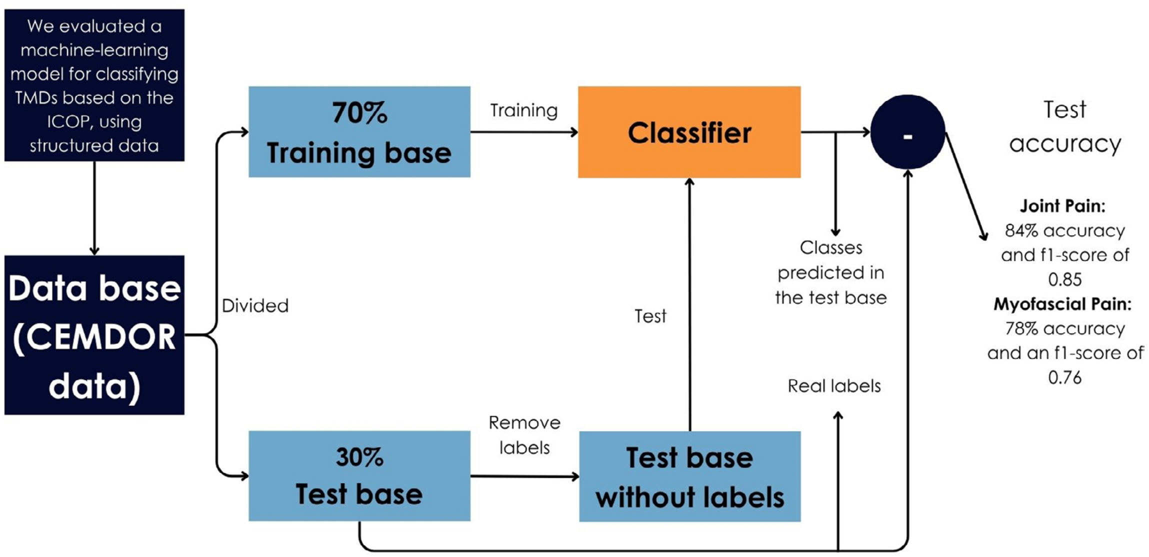
Fernanda Pretto ZATT[1] | João Victor Cunha CORDEIRO[1] | Lauren BOHNER[1] | Beatriz Dulcineia Mendes de SOUZA[1] | Victor Emanoel Armini CALDAS[2] | Ricardo Armini CALDAS[1]

[1]Universidade Federal de Santa Catarina (UFSC), Departamento de Odontologia, Florianópolis, Brasil.
[2]Independent Researcher, Amsterdam, Netherlands.

Abstract

Background: the escalating influx of patients with temporomandibular disorders and the challenges associated with accurate diagnosis by non-specialized dental practitioners underscore the integration of artificial intelligence into the diagnostic process of temporomandibular disorders (TMD) as a potential solution to mitigate diagnostic disparities associated with this condition. Objectives: In this study, we evaluated a machine-learning model for classifying TMDs based on the International Classification of Orofacial Pain, using structured data. Methodology: Model construction was performed by the exploration of a dataset comprising patient records from the repository of the Multidisciplinary Orofacial Pain Center (CEMDOR) affiliated with the Federal University of Santa Catarina. Diagnoses of TMD were categorized following the principles established by the International Classification of Orofacial Pain (ICOP-1). Two independent experiments were conducted using the decision tree technique to classify muscular or articular conditions. Both experiments uniformly adopted identical metrics to assess the developed model's performance and efficacy. Results: The classification model for joint pain showed a relevant potential for general practitioners, presenting 84% accuracy and f1-score of 0.85. Thus, myofascial pain was classified with 78% accuracy and an f1-score of 0.76. Both models used from 2 to 5 clinical variables to classify orofacial pain. Conclusion: The use of decision tree-based machine learning holds significant support potential for TMD classification.

Keywords: Artificial intelligence. Machine learning. Facial pain. Diagnosis.



The utilization of decision tree-based machine learning holds significant potential to support in the classification of TMD.

*The article is a study produced from a master dissertation available from:https://repositorio.ufsc.br/xmlui/handle/123456789/251470

Corresponding address:
Fernanda Pretto Zatt - Universidade Federal de Santa Catarina - Departamento de Odontologia - Rua Delfino Conti, 1240 - 88040-535 - Florianópolis - SC - Brasil.
Phone: +55 48 3721 4952
e-mail: ferdpretto@gmail.com

Received: July 8, 2024
Revised: September 14, 2024
Accepted: September 18, 2024

Editor: Linda Wang
Associate Editor: Paulo César Conti

Introduction

Temporomandibular disorders (TMD) encompass a diverse array of conditions that affect the masticatory muscles, the temporomandibular joint (TMJ), and their associated structures.[1] These conditions may include a broad spectrum of signs and symptoms, ranging from an isolated condition to the involvement of multiple systems. This complexity renders the diagnosis of TMDs a challenging task with a significant potential for diagnostic errors, particularly for non-specialist professionals in the field.[2–6] The classification of TMDs is typically grounded in international consensus among experts[7,8] and entails a comprehensive patient assessment that takes many factors into account.

The International Classification of Orofacial Pain (ICOP-1) is the first comprehensive classification system solely dedicated to orofacial pain. This classification was modeled after the structure of the International Classification of Headache Disorders (ICHD), which is widely accepted and globally used by medical practitioners and researchers.[7,9] However, these conditions are often underdiagnosed or receive limited attention in oral healthcare, which can be observed in the disparity between the estimated amount of required treatment and the treatment actually performed.[10]

The literature provides evidence of deficiencies in dentists' knowledge on the management of patients with TMD.[11–13] Furthermore, studies have demonstrated that a substantial proportion of such professionals lack the necessary experience and skills to perform diagnostic and therapeutic procedures related to TMD.[14] One potential explanation for this situation is the inadequate training received during undergraduate courses, which can impact the ability of young and inexperienced dentists to identify TMD.[10,14] In the current era of dental education, however, there has been a shift toward digitalization and the integration of new technologies to enhance the learning process.[15,16] These advances, such as digital simulations and artificial intelligence-based tools, have the potential to significantly improve diagnostic capabilities and clinical training for TMD. Nevertheless, despite these benefits, the implementation of such technologies can also introduce challenges, as they may increase stress and dissatisfaction among students due to the steep learning curve of these tools.[17]

In recent years, various screening tools and clinical examinations for detecting temporomandibular disorders (TMD) have been developed and used to address gaps in the diagnosis of these conditions. These tools should be relatively concise, use plain language, and be easy to implement.[18] Furthermore, efforts to simplify and adapt assessment systems to make them more feasible for general practitioners are commendable. Nevertheless, it is crucial that each clinician becomes familiar with the recently proposed TMD assessment tools and incorporates them into their practice.[18]

In the other hand, the application of artificial intelligence (AI) in TMD diagnosis holds the potential to provide support for non-specialist professionals, resulting in a more effective dental practice,[2–5] given its capacity to enable computers to act like humans and make independent decisions or requiring minimal human intervention. AI use is not confined solely to academic research but extends into the commercial domain, encompassing various areas within dentistry.[19] Its applicability can span from handling dental emergencies to the differential diagnosis of oral lesions, interpretation of radiographs, analysis of facial growth during orthodontic treatments, planning and simulation of ideal restorations for specific patients, and has demonstrated promising outcomes in detecting TMD.[2,5,20-23]

Among the various available AI methods, the decision tree algorithm via machine learning functions similarly to human reasoning in decision-making. Decision trees constitute a group of classifiers with comprehensible criteria and enabling their logic to be evaluated by an expert.[24] From the perspective of knowledge discovery, the capacity to monitor and assess each step of the decision-making process emerges as a significant factor for instilling confidence in auxiliary diagnostic methods within the healthcare domain.[25] Thus, this study's objective was to develop and assess a machine-learning model with the capability to classify TMD based on the ICOP-1.

Methodology

The analyses were conducted using the clinical records repository of the Multidisciplinary Orofacial Pain Center (CEMDOR) at the Federal University of Santa Catarina, Brazil. All diagnoses were either provided or reviewed by a specialist. This center holds a distinguished reference position in the field of orofacial pain, catering to patients referred from a wide range of public and private institutions. An operator performed the initial screening of the patient records and subsequently transcribed the relevant information into a comprehensive database. The project was approved by the institutional ethics committee (CAAE: 60532822.8.0000.0121).

Inclusion criteria were patients aged at least 18 years, with a diagnosis of joint or muscle TMD and complete clinical documentation. Patients with incomplete treatment documentation, as well as records in which subjective parts were described in continuous text, were excluded from the sample. The inclusion criteria generated a convenience sample of 122 medical records, of which the diagnoses are outlined in Table 1.

The execution of the machine learning phase was carried out using the decision tree method. For this purpose, Python programming language (Python Software Foundation, Delaware, USA) was selected, along with the libraries NumPy, Pandas, Scikit-learn, Graphviz, Matplotlib, and Seaborn. These tools are code libraries that offer an array of functionalities for program development. Initially, the samples were randomly divided into proportions of 70% and 30% for the construction of training and testing models, respectively. To enhance model performance, three parameters were pre-adjusted: the function to measure the quality of a split, the maximum depth of the tree, and the minimum number of samples required to split an internal node. The patients' record data in CSV file format were loaded into a Pandas DataFrame object. Two distinct decision tree models were constructed: one for orofacial myofascial pain and another for temporomandibular joint (TMJ) pain. The following steps and descriptive analyses were conducted independently for each of the models.

Attempts were undertaken using all combinations of two functions to measure the quality of a split (Gini Index and Entropy): four potential maximum depths of the tree (2, 3, 4, and 5), and six potential minimum number of samples required to split an internal node (5, 10, 15, 20, 25, and 30). The combinations that yielded the highest accuracies for the training data (78% orofacial myofascial pain and 84% TMJ pain) were subsequently chosen for further analysis.

Results

Two independent experiments were carried out to classify myofascial pain and TMJ pain. Both experiments used the same metrics to evaluate the models performance.

Orofacial myofascial pain

Table 2 shows the metrics used to evaluate the classification model for orofacial myofascial pain. These metrics provide information on the model performance for each class individually, as well as aggregate metrics that consider the model overall performance.

The decision tree employed data from specific medical records for decision-making, using five characteristics out of a total of 54. The most relevant characteristic was the onset of pain (0.514), followed by location (0.304), maximum opening, the visual analogue scale of pain, and pain by palpation of the posterior TMJ. Other characteristics supplied to the model did not contribute significantly to predicting the outcome.

Temporomandibular joint pain

Table 3 shows the metrics used to evaluate the classification model for temporomandibular joint (TMJ) pain. These metrics provide information on the model performance for each class individually, as well as aggregate metrics that consider the model overall performance.

Table 1- Diagnoses from medical records

Diagnostic	Amount
Myofascial orofacial pain	
Primary myofascial orofacial pain	31
Acute primary myofascial orofacial pain	11
Chronic primary myofascial orofacial pain	80
Temporomandibular joint (TMJ) pain	
Primary temporomandibular joint pain	72
Secondary temporomandibular joint pain	34

Table 2- Performance metrics of the decision tree model for orofacial myofascial pain

	Precision	Recall	F1-score	Support
Primary myofascial orofacial pain	0.83	0.50	0.62	10
Acute primary myofascial orofacial pain	0.50	0.33	0.40	3
Chronic primary myofascial orofacial pain	0.79	0.96	0.87	24
accuracy	0.78			37
macro avg	0.71	0.60	0.63	37
weighted avg	0.78	0.78	0.76	37

Table 3- Performance metrics of the decision tree model for temporomandibular joint pain

	Precision	Recall	F1-score	Support
Primary TMJ joint pain	0.90	0.86	0.88	22
Secondary TMJ joint pain	0.73	0.80	0.76	10
accuracy	0.84			32
macro avg	0.82	0.83	0.82	32
weighted avg	0.85	0.84	0.85	32

The decision tree employed data from specific medical records for decision-making in TMJ pain cases, using only two characteristics out of a total of 54. The most relevant characteristic was the presence of a "clicking" effect at the beginning of the mouth opening (0.946), followed by pain by palpation of the temporal muscle (0.005).

Discussion

This study describes a machine-learning approach based on a decision tree for the detection of orofacial myofascial pain and TMJ pain, with diagnoses based on the ICOP-1. Two different decision trees were generated, one for each condition assessed. Results show that the model effectively detected both conditions with an accuracy of over 75%. Therefore, our results suggest that AI can play a valuable complementary role in the existing diagnostic tools. The model was developed using the medical records of a reference center for orofacial pain – CEMDOR. Thus, every patient had diagnosis of some level of pain, whether muscular and articular or just muscular. Therefore, after screening all the medical records, we obtained a convenience sample of 122 complete records with diagnoses reviewed by a specialist in TMD and orofacial pain.

In the case of orofacial myofascial pain, an accuracy of 78% was obtained for classifying primary orofacial myofascial pain, acute primary orofacial myofascial pain, and chronic primary orofacial myofascial pain. These data are similar to previous studies on the use of AI to detect painful conditions of the face.[5,6] Our model obtained a good overall performance in classifying the proposed conditions, similar to previous studies using AI.[3–6] In the analysis of muscle conditions, underdiagnoses were observed for the class of acute primary pain, in which one subject was classified as chronic primary and another only as primary pain. The above-described difficulty in correct classification may be due to the small number of medical records provided to the model with this type of diagnosis when compared to the other classes. The identification of TMD is a long process based on medical history and complementary exams, with evidence suggesting that acute pain is not early detected nor is a condition diagnosed at an early stage.[10]

For the detection of TMJ pain, an accuracy of 84% was found for classifying primary TMJ pain and secondary TMJ pain. Models created to detect joint TMD using imaging tests achieved similar accuracy to our study, with 78% accuracy when using CT scans – similarly to that achieved by specialist radiologists.[4] Studies using magnetic resonance imaging have achieved 89% accuracy.[3] Both studies using data from complementary imaging exams and clinical data from medical records achieved relevant diagnostic accuracy values; however, no study managed to classify all diagnoses correctly. Since the association of both analyses is essential for the diagnosis of 100% of cases, it should be considered for improvement in AI-assisted diagnostic models.

In addition to the presented data, it was possible to observe the small number of variables used to diagnose orofacial pain conditions. The complexity and excessive data volume in medical records for pain diagnosis make clinical interpretation difficult, especially for professionals without expertise in the area. From a clinical point of view, all the questions in the questionnaire are relevant and contribute to designing an appropriate and personalized treatment for the patient. However, the model used only a few characteristics out of several to predict a condition and, when provided with an adequate amount of data, obtained correct diagnoses.

Extensive diagnostic instruments are difficult to manage, both in clinical settings and epidemiological research due to the time they consume and the fatigue generated in both patients and professionals, affecting the answers obtained and the way they are recorded. Despite the diagnosis complexity, it is important that the instrument is simple and contains only relevant questions.[26] However, it is important to note that this study did not cover all orofacial pain diagnoses. To define relevant variables until the last stage of each possible diagnosis, it is necessary to carry out a more comprehensive analysis, making it possible to determine whether other characteristics are relevant to the diagnosis of the conditions in question.

To improve our model, it is necessary to increase the number of patients and balance the distribution between diagnoses. One of the limitations of the decision tree is its susceptibility to underfitting and overfitting, especially when data sample is small.[25,27] Underfitting occurs when the model does not fit the training data because it fails to learn important patterns in the data. This problem is identified when the model performs poorly on the training and test data. Overfitting occurs when the model overfits the training data, capturing even the noise present in the data. In this case, the model loses its ability to generalize to new data. Overfitting can be identified when a model performs excellently on training data but poorly on test data.[28] This compromises the classification capacity and robustness of the model.[27]

Other important aspect observed in this study is the identification of seemingly unconventional characteristics by the machine learning model when distinguishing between orofacial myofascial pain and temporomandibular joint pain. These associations, such as the link between posterior TMJ palpation and myofascial pain, may appear counterintuitive from a clinical perspective. This reflects a common limitation in AI models, which, when trained, identify statistical patterns that do not always align with

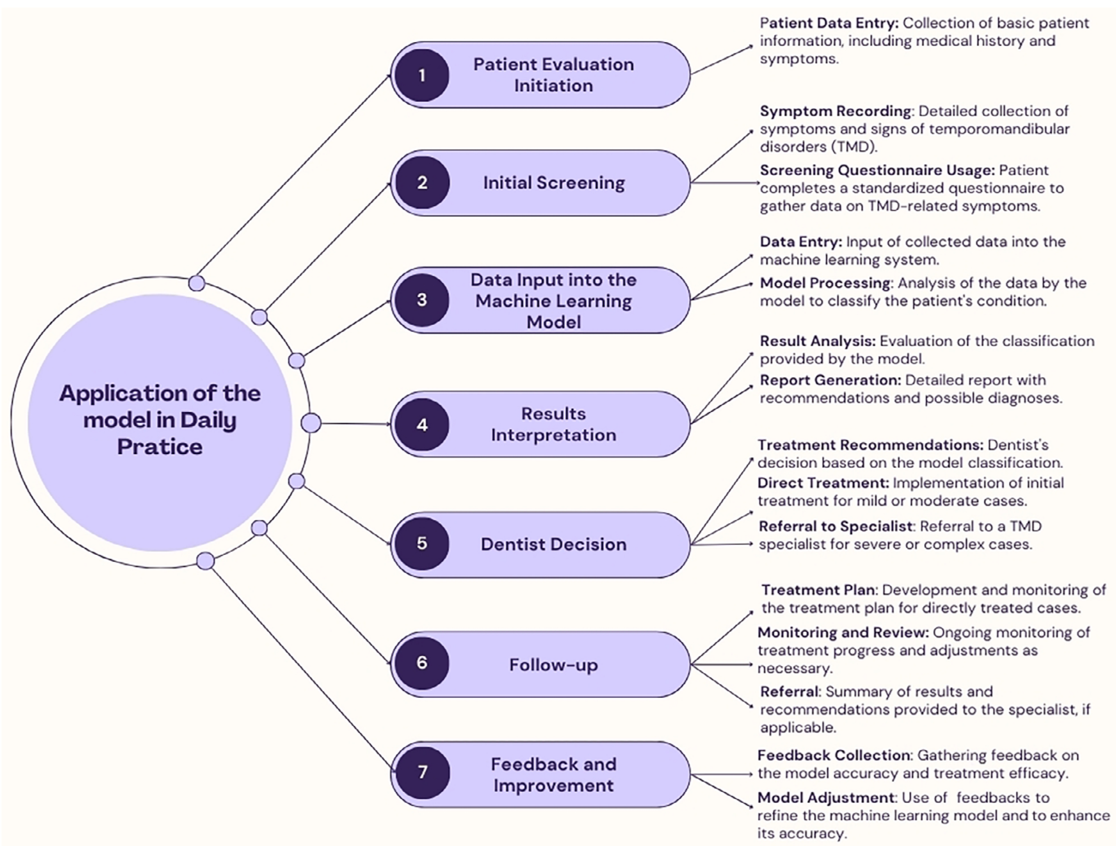


Figure 1- Flowchart of machine learning model application in daily dental practice to assist dentists in decision-making

established clinical practice. The integration of continuous feedback from healthcare professionals and the adjustment of model parameters can help in reducing these discrepancies and improving the results applicability in clinical settings.

In addition to these limitations, the use of artificial intelligence (AI) in TMD assessment also presents challenges. While AI offers great potential for improving diagnostic accuracy, it requires large, high-quality datasets to train robust models. Furthermore, the clinical integration of AI tools requires extensive validation and testing to ensure reliability in diverse patient populations. Another limitation is the potential for AI models to introduce bias if the training data is not sufficiently representative of the broader population. These factors must be considered when interpreting the results of AI-based models in TMD assessment, and future studies should focus on addressing these challenges to enhance the accuracy and applicability of such technologies in clinical practice.

Although diagnostic methods aided by artificial intelligence (AI) are less relevant when used by specialists, they are highly valuable for non-specialist clinicians.[5] Figure 1 illustrates how the model developed in this study can be applied in daily dental practice to assist dentists in decision-making. This flowchart provides a clear framework for integrating the model into routine clinical workflows, helping dentists efficiently assess and manage temporomandibular disorders. To this end, AI tools must be developed with scientific rigor, validated through clinical studies, and used as a complementary tool in the clinical judgment of health professionals. The application of AI in dentistry represents great potential to drive advances and benefits in the field. It is a powerful tool that complements and assists the work of dental surgeons, providing faster and more accurate diagnoses, as well as a better overall patient experience. In addition, AI can help analyze large amounts of dental data for a better understanding of oral health patterns, trends, and predictions, helping to develop more effective prevention and treatment strategies.

Conclusions

It can be concluded that:

The use of decision trees by machine learning presents relevant potential to aid in the clinical diagnosis of TMD;

The implementation of TMD clinical diagnosis methods aided by AI should be widely adopted in the clinical routine;

Non-specialist professionals in the field of TMD and orofacial pain can benefit from AI-assisted clinical diagnosis.

Acknowledgments

The authors would like to thank the CAPES (Coordination for the Improvement of Higher Education Personnel) agencies and the Federal University of Santa Catarina (UFSC).

Conflict of interest

The authors declare no conflict of interest.

Data availability

The datasets generated and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors´contributions

Zatt, Fernanda Pretto: Conceptualization (Equal); Data curation (Equal); Funding acquisition (Equal); Investigation (Equal); Writing - original draft (Equal). **Cordeiro, João Victor Cunha**: Conceptualization (Equal); Data curation (Equal); Investigation (Equal); Writing - original draft (Equal). **Bohner, Lauren**: Conceptualization (Equal); Data curation (Equal); Supervision (Equal); Writing - review & editing (Equal). Souza, **Beatriz Dulcinéia Mendes**: Conceptualization (Equal); Project administration (Equal); Validation (Equal); Visualization (Equal). **Caldas, Victor Emanoel Armini:** Conceptualization (Equal); Methodology (Equal); Validation (Equal); Visualization (Equal). **Caldas, Ricardo Armini**: Conceptualization (Equal); Formal analysis (Equal); Methodology (Equal); Supervision (Equal); Validation (Equal); Writing - review & editing (Equal).

References

1- Leeuw R, Klasser GD. Orofacial pain: guidelines for assessment, diagnosis, and management. Stomatology Edu J. 2015;2(2):173; doi: 10.25241/stomaeduj.2015.2(2).bookreview.6.

2- Farook TH, Jamayet N Bin, Abdullah JY, et al. Machine learning and intelligent diagnostics in dental and orofacial pain management: a systematic review. Pain Res Manag. 2021;2021:6659133 doi: 10.1155/2021/6659133

3- Orhan K, Driesen L, Shujaat S, Jacobs R, Chai X. Development and validation of a magnetic resonance imaging-based machine learning model for tmj pathologies. Biomed Res Int. 2021;2021: 6656773 doi: 10.1155/2021/6656773

4- Choi E, Kim D, Lee JY, Park HK. Artificial intelligence in detecting temporomandibular joint osteoarthritis on orthopantomogram. Sci Rep. 2021;11(1): 10246 doi: 10.1038/s41598-021-89742-y

5- Reda B, Contardo L, Prenassi M, Guerra E, Derchi G, Marceglia S. Artificial intelligence to support early diagnosis of temporomandibular disorders: a preliminary case study. J Oral Rehabil. 2023;50(1):31-8. doi: 10.1111/joor.13383

6- Calil BC, Cunha DV, Vieira MF, Oliveira Andrade A, Furtado DA, et al. Identification of arthropathy and myopathy of the temporomandibular syndrome by biomechanical facial features. Biomed Eng Online. 2020;19(1):22. doi: 10.1186/s12938-020-00764-5

7- International Classification of Orofacial Pain, 1st edition (ICOP). Cephalalgia. 2020;40(2):129-21. doi: 10.1177/0333102419893823

8- Schiffman E, Ohrbach R, Truelove E, Look J, Anderson G, Goulet JP, et al. Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) for clinical and research applications: recommendations of the International RDC/TMD Consortium Network * and Orofacial Pain Special Interest Group. J Oral Facial Pain Headache 2014;28(1):6-27. doi: 10.11607/jop.1151.

9- Pigg M, Nixdorf DR, Law AS, Renton T, Sharav Y, Baad-Hansen L, et al. New International Classification of Orofacial Pain: What Is in It For Endodontists? J Endod 2021;47(3):345-57. doi: 10.1016/j.joen.2020.12.002

10- Lövgren A, Visscher CM, Häggman-Henrikson B, Lobbezoo F, Marklund S, Wänman A. Validity of three screening questions (3Q/TMD) in relation to the DC/TMD. J Oral Rehabil. 2016;43(10):729-36. doi: 10.1111/joor.12428

11- Aggarwal VR, Joughin A, Zakrzewska JM, Crawford FJ, Tickle M. Dentists' and specialists' knowledge of chronic orofacial pain: results from a continuing professional development survey. Primary Dental Care. 2011;18(1):41-4. doi: 10.1308/135576111794065838

12- Hadlaq E, Khan H, Mubayrik A, Almuflehi NS, Mawardi H. Dentists' knowledge of chronic orofacial pain. Niger J Clin Pract. 2019;22(10):1365. doi: 10.4103/njcp.njcp_110_19

13- Al-Huraishi HA, Meisha DE, Algheriri WA, Alasmari WF, Alsuhaim AS, Al-Khotani AA. Newly graduated dentists' knowledge of temporomandibular disorders compared to specialists in Saudi Arabia. BMC Oral Health 2020;20(1):272. doi: 10.1186/s12903-020-01259-4

14- Reissmann DR, Behn A, Schierz O, List T, Heydecke G. Impact of dentists' years since graduation on management of temporomandibular disorders. Clin Oral Investig. 2015;19(9):2327-36. doi: 10.1007/s00784-015-1459-7

15- Erdilek D, Gümüştaş B, Güray Efes B. Digitalization era of dental education: a systematic review. Dent Med Probl. 2023;60(3):513-25. doi: 10.17219/dmp/156804

16- Delgado Castillo S, Miguel Soto S, Atoche-Socola K, Arriola-Guillén LE. Revolution in modern teaching in dentistry since the appearance of the COVID-19 pandemic: a review. Dent Med Probl. 2022;59(1):137-41. doi: 10.17219/dmp/141522

17- Ertürk Avunduk AT, Delikan E. Satisfaction and stress levels of dentistry students relating to distance education. Dent Med Probl. 2021;58(3):291-8. doi: 10.17219/dmp/135318

18- Emodi-Perlman A, Eli I. Temporomandibular disorders and bruxism – up-to-date assessment and screening tools the general dentist should be aware of. Dent Med Probl. 2024;61(2):169-71. doi: 10.17219/dmp/175582

19- MacHoy ME, Szyszka-Sommerfeld L, Vegh A, Gedrange T, Woźniak K. The ways of using machine learning in dentistry. Adv Clin Exp Med. 2020;29(3):375-84. doi: 10.17219/acem/115083

20- Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science. 2015;349(6245):255-60. doi: 10.1126/science.aaa8415

21- Schwendicke F, Samek W, Krois J. Artificial intelligence in dentistry: chances and challenges. J Dent Res. 2020;99(7):769-74. doi: 10.1177/0022034520915714

22- Kose C Jr, Oliveira D, Pereira PN, Rocha MG. Using artificial intelligence to predict the final color of leucite-reinforced ceramic restorations. J Esth Restor Dent. 2023;35(1):105-15. doi: 10.1111/jerd.13007

23- Carrillo-Perez F, Pecho OE, Morales JC, Paravina RD, Della Bona A, Ghinea R. Applications of artificial intelligence in dentistry: a comprehensive review. J Esthet Rest Dent. 2022;34(1):259-80. doi: 10.1111/jerd.12844

24- Loyola-Gonzalez O. Black-box *vs*. white-box: understanding their advantages and weaknesses from a practical point of view. IEEE Access. 2019;7:154096-113. doi: 10.1109/ACCESS.2019.2949286

25- Stiglic G, Kocbek S, Pernek I, Kokol P. Comprehensive decision tree models in bioinformatics. PLoS One. 2012;7(3):e33812. doi: 10.1371/journal.pone.0033812

26- Firmino RT, Granville-Garcia AF, Bendo CB, Ferreira FM, Ortiz FR, Mcgrath CP. Development and validation of a short form of the BOHLAT-P. Braz Oral Res. 2022;36:e074. doi: 10.1590/1807-3107bor-2022.vol36.0074

27- Li Y, Sun G, Zhu Y. Data imbalance problem in text classification. In: 2010 Third International Symposium on Information Processing IEEE; 2010. p. 301-305. doi: 10.1109/ISIP.2010.47

28- Hernández VAS, Monroy R, Medina-Pérez MA, Loyola-Gonçalves O, Herrera F. A practical tutorial for decision tree induction. ACM Comput Surv. 2022;54(1):1-38. doi: 10.1145/3429739