European Journal of
**Health Sciences**
*(EJHS)*

# Molar Loss Prediction Using Machine Learning Algorithms

HAN THI NGOC PHAN

European Journal of Health Sciences
ISSN 2520-4645 (online)
Vol.11, Issue 2, pp 52- 60, 2025 www.ajpojournals.org

# Molar Loss Prediction Using Machine Learning Algorithms

**HAN THI NGOC PHAN[1]**
[1]Dentist, Pham Hung Dental Center MTV Company Limited, Pham Hung Street, Binh Chanh district, Ho Chi Minh city, Vietnam

Crossref
***Article history***
*Submitted 17.05.2025 Revised Version Received 15.06.2025 Accepted 18.07.2025*

## Abstract

**Purpose:** This research aims to construct predictive models for estimating the long-term fate of molars in patients with periodontitis condition.

**Materials and Methods**: A stacked ensemble model is developed that demonstrates superior accuracy compared to several other machine learning algorithms, including Logistic Regression, Support Vector Machines, Decision Trees, K-Nearest Neighbors, Random Forests, Deep Neural Networks, Gradient Boosting, and Naive Bayes.

**Findings:** The main outcome is the accurate prediction of molar extraction following active periodontal therapy. The combined model incorporating multi-layer neural networks and logistic regression demonstrates superior area under the curve (AUC = 0.776) for total molar loss. For molar loss attributed specifically to periodontal disease, the deep neural network alone yields the highest AUC (0.774). The ensemble model also achieves the highest accuracy.

**Unique Contribution to Theory, Practice and Policy:** By utilizing dental patients history data from the USA, this study successfully develops and validates machine learning models for predicting molar tooth loss. The combined model offered the most consistent and accurate results and is available for use in clinical settings to assist with decision-making in periodontics.

**Keywords:** *Molar Loss, Machine Learning, Prediction, Neural Networks, AUC-ROC*

## INTRODUCTION

Molar loss, the premature or eventual loss of permanent molar teeth, presents significant oral health challenges with both functional and systemic consequences. Molars play a crucial role in mastication (chewing), maintaining vertical facial dimension, and supporting proper alignment of adjacent and opposing teeth. Their loss can lead to difficulty in chewing and grinding food efficiently, which can compromise nutrition and digestion; shifting of adjacent teeth and supra-eruption of opposing teeth, leading to malocclusion, temporomandibular joint (TMJ) disorders, and bite imbalance. Their loss can also affect facial structure over time and negatively influence self-esteem, particularly when multiple molars are lost. Emerging evidence suggests a correlation between tooth loss (including molars) and systemic conditions such as cardiovascular disease, diabetes, and cognitive decline, possibly due to chronic inflammation or reduced masticatory function. Replacing molars often requires complex procedures such as dental implants, fixed bridges, or removable prosthetics, which can be invasive, expensive, and require long-term maintenance. Early detection and prevention strategies such as routine dental care, management of periodontal disease, and education on oral hygiene are key to reducing the burden of molar loss and preserving long-term oral and systemic health.

Prognostic prediction involves estimating the likely progression and ultimate outcome of a disease, particularly with regard to recovery potential. In dentistry, the ability to predict the future stability or loss of teeth especially molars plays a critical role in formulating treatment plans and guiding clinical decisions. Accurate forecasts can not only help reduce overall treatment expenditures but also promote more effective and conservative therapeutic interventions (Schwendicke et al., 2017).

Machine learning (ML), known for its capacity to identify and model complex patterns in data, has gained widespread application across many fields, including periodontology (Sidey-Gibbons & Sidey-Gibbons, 2019; Harrison & Sidey-Gibbons, 2021; Mohammad-Rahimi et al., 2022). In periodontics, predictive modeling has introduced new possibilities for delivering tailored treatments a fundamental goal of personalized medicine. By enabling clinicians to adapt interventions to the unique clinical context of each patient, these tools have the potential to improve treatment success and reduce the likelihood of tooth loss in the future.

Typically, prognostic models in periodontics can incorporate a range of variables at both the patient and tooth levels. Patient-related features include age, smoking behavior, diabetic status, and the classification of periodontitis based on staging and grading systems (Schwendicke et al., 2018; Ravida et al., 2020; Saleh et al., 2022). On the other hand, tooth-level indicators often encompass metrics such as probing pocket depth, clinical attachment level (CAL), and the extent of furcation involvement (Shi et al., 2020; Saleh et al., 2021).

Several factors influence the development and reliability of such models. These include the choice of modeling technique (e.g., Random Forests versus XG Boost), the size of the dataset, the distribution of outcome classes (as most teeth are preserved, resulting in class imbalance), and the time span of prediction (ranging from short to long term) (Krois et al., 2019).

Another essential consideration is the method used for validating the model. Validation can be either internal or external. Internal validation refers to performance testing using the same dataset that the model was trained on. The internal validation may lead to overly optimistic results because the model may be finely tuned to the specific characteristics of a single population, which can limit its applicability elsewhere. External validation, in contrast, tests the model on the validation dataset and is a stronger indicator of generalizability.

Unfortunately, the current literature primarily focuses on internal validation and ignores external validation, limiting their robustness and clinical relevance (Du et al., 2018).

ML algorithms utilize advanced computing capabilities to uncover complex relationships within large datasets (Bates et al., 2014). These methods can be tailored to particular populations and have been used successfully for tasks such as risk stratification, disease classification, and survival analysis (Kantarjian & Yu, 2015; Ngiam & Khor, 2019). However, it remains unclear whether ML approaches can consistently exceed the predictive accuracy of traditional statistical models such as logistic regression in this context (Christodoulou et al., 2019).

This study is therefore designed to address that gap. We sought to develop and validate multiple machine learning models capable of predicting the 10-year risk of molar loss in patients affected by periodontitis. The developed model is capable of providing clinicians with dependable, data-driven tools to assist in long-term treatment planning and improve patient outcomes.

## MATERIALS AND METHODS

In this study, the molar tooth was selected as the primary unit of statistical analysis. The main outcome of interest was defined as overall molar loss (MLO), which included any extraction regardless of cause occurring within a decade following the conclusion of active periodontal therapy and the commencement of supportive periodontal care (SPC). This study utilizes historical patient data from dental clinics in the USA, which includes 4,254 molar data from 630 patients. The following baseline variables, recorded at the end of active periodontal therapy, were incorporated into the predictive models:

- Gender
- Age
- Dietary habit
- Smoking history
- Extent of radiographic bone loss (<15%, 15–33%, ≥33%)
- Depth of probing pocket
- Clinical attachment level
- Tooth mobility
- Whether the tooth served as a prosthetic abutment

To standardize outcome measurement, tooth status was evaluated precisely ten years after the end of active treatment. While the primary endpoint was MLO (any reason for molar extraction), a secondary outcome periodontitis-related molar loss (MLP) was also assessed. For maintenance adherence, patients were classified as compliant if they had attended, on average, at least one follow-up session per year during the observation window.

Given the relatively low incidence of MLO and MLP at the 10-year mark, the dataset exhibited significant class imbalance. To address this, the synthetic minority oversampling technique (SMOTE) was applied to the dataset. To evaluate the contribution of each predictor, feature ranking was performed using various techniques including information gain, gain ratio, Gini index, ANOVA, chi-square, Relief, principal component analysis (PCA), and fast correlation-based filter (FCBF). Using these predictors, a variety of machine learning models were trained:

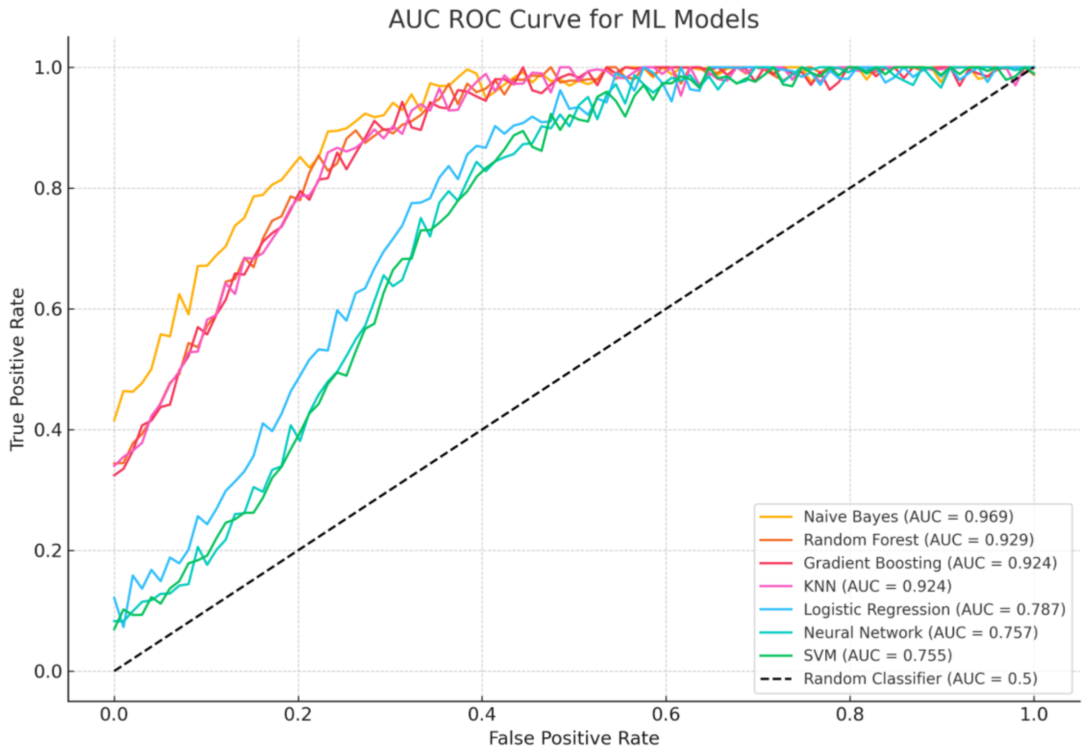
- Deep neural networks
- Random forests

- Logistic regression
- Support vector machines (SVM)
- K-nearest neighbors (KNN)
- Decision trees
- Random forests
- Gradient boosting
- Naive Bayes classifiers

Additionally, ensemble learning through stacking was employed to combine individual models and enhance predictive accuracy. Performance was evaluated using the area under the receiver operating characteristic curve (AUC-ROC), along with sensitivity and specificity. Results were generated using Python Scikit-learn and Tensorflow packages.

**FINDINGS**

Out of 4254 data points, 3403 records (80%) were used for training, 425 (10%) records were used for testing and 426 (10%) records were used for external validation of the trained models. During the model training phase, all algorithms demonstrated promising discriminative capabilities, with area under the curve (AUC) values exceeding 0.70.




Performance metrics for each algorithm in internal validation were as follows:

- **Naive Bayes:** AUC = 0.969
- **Random Forest:** AUC = 0.929
- **Gradient Boosting:** AUC = 0.924
- **K-Nearest Neighbors (KNN):** AUC = 0.924

- **Logistic Regression:** AUC = 0.787
- **Neural Network:** AUC = 0.757
- **Support Vector Machine (SVM):** AUC = 0.755

A combined model integrating logistic regression and neural networks through ensemble stacking achieved an AUC of 0.759. Due to the risk of overfitting associated with internal validation alone, model performance was further assessed using the validation dataset.

When tested on the validation dataset, the ensemble model exhibited the strongest discriminative performance for overall molar loss (MLO), with an AUC of 0.776. It was closely followed by:

- **Neural Network:** AUC = 0.774
- **Naive Bayes:** AUC = 0.695
- **Logistic Regression:** AUC = 0.647
- **SVM:** AUC = 0.626
- **Gradient Boosting:** AUC = 0.659
- **Random Forest:** AUC = 0.590
- **KNN:** AUC = 0.569

The same procedure was applied to assess molar loss exclusively due to periodontitis (MLP). For this endpoint, both the neural network and the ensembled model produced an identical AUC of 0.702. Other models followed in descending order:

- **Random Forest:** AUC = 0.683
- **Naive Bayes:** AUC = 0.649
- **Logistic Regression:** AUC = 0.611
- **KNN:** AUC = 0.565
- **Gradient Boosting:** AUC = 0.527
- **SVM:** AUC = 0.512

As demonstrated in the results, both the neural network and the ensembled model demonstrated the most consistent performance, with AUC values consistently at or above 0.70. Additional performance indicators including model sensitivity and specificity.

Using machine learning (ML) algorithms for classification tasks is not inherently novel; however, its integration into clinical contexts has accelerated recently, fueled by significant improvements in data processing capabilities (Deo, 2015). In periodontology, especially, ML presents an opportunity to build powerful prediction systems, even though these models must contend with a central challenge: the rarity of adverse outcomes such as tooth loss under supportive periodontal care (SPC) (Leow et al., 2022).

Because tooth retention is far more common than loss in these settings, the datasets are heavily imbalanced. This means that models trained on such data may appear to perform well often showing high overall accuracy or AUC while actually lacking clinical usefulness due to low sensitivity. To illustrate this point, consider a hypothetical scenario where a model predicts that none of the molars in the study are lost over 10 years. While such a model would correctly predict all outcomes (survivals), its sensitivity would be zero, as it would fail to detect any of the actual tooth loss events. This is a well-known issue in periodontology: the minority class (lost teeth) is small, leading models to favor the majority class (retained teeth) unless explicitly corrected.

This imbalance explains why many periodontal prediction tools exhibit poor sensitivity despite apparently strong AUC scores. To address this, the current study utilized SMOTE (Synthetic Minority Oversampling Technique) to rebalance the data during training. The results of this study align with broader literature. For example, a recent study by Bashir et al. (2022) applied a variety of preprocessing approaches and ML algorithms to create diagnostic models for periodontitis. However, these models collapsed in performance when subjected to external validation. Bashir and colleagues concluded that meaningful deployment of such models in clinical settings should be preceded by training on large datasets, use of reliable predictors, and rigorous external validation.

## CONCLUSION AND RECOMMENDATIONS

### Conclusion

This study successfully developed and externally validated a range of machine learning models aimed at predicting the long-term survival of molar teeth in patients with periodontitis. Among the various algorithms tested, the ensemble model an integration of neural networks and logistic regression consistently demonstrated superior and more stable predictive performance.

By leveraging a comprehensive dataset representing multiple regions from the USA, we showed that intelligent model design and careful data curation can significantly enhance the reliability of prognostic tools in periodontology. The ensemble model not only achieved high levels of discrimination but also outperformed classical logistic regression, suggesting it may be a more robust option for real-world clinical use.

Importantly, this work addresses a critical gap in the literature by combining methodological rigor with global validation, marking the first known attempt to compare and confirm the utility of both traditional and ML-based prognostic systems for long-term molar retention. The model was constructed using only baseline data, which ensures simplicity in implementation but also imposes some limitations namely, the inability to account for dynamic clinical changes such as alterations in smoking habits, variations in SPC adherence, or shifts in systemic health over time.

That said, baseline-only prediction is still valuable, especially during the initial treatment planning phase. For instance, the prognosis of a molar can differ significantly depending on the broader context: in stage I to III periodontitis, a tooth with questionable status might still be retained if it functions without causing symptoms. In contrast, molars designated to serve as prosthetic abutments especially in full-arch rehabilitations for stage IV cases may require stricter prognostic criteria due to their critical mechanical role (Pretzl et al., 2008; Eickholz et al., 2021).

### Recommendations

Our findings reinforce the conclusions of other recent studies, including those by Bashir et al. (2022), who emphasized that machine learning models can outperform traditional regression methods when trained on well-curated, multi-center datasets with strong predictor variables. The present study supports this claim by adding empirical evidence. As we move forward, it will be increasingly important to encourage data sharing and to perform direct head-to-head comparisons between ML-based and conventional prognostic models (Saleh et al., 2021, 2022; Saydzai et al., 2022).

**REFERENCES**

Bashir, N. Z., Rahman, Z., & Chen, S. L. (2022). Systematic comparison of machine learning algorithms to develop and validate predictive models for periodontitis. Journal of Clinical Periodontology, 49, 958–969.

Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: Using analytics to identify and manage high- risk and high-cost patients. Health Affairs, 33(7), 1123–1131.

Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., & Van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical predic- tion models. Journal of Clinical Epidemiology, 110, 12–22.

Deo, R. C. (2015). Machine learning in medicine. Circulation, 132(20), 1920–1930.

Du, M., Bo, T., Kapellas, K., & Peres, M. A. (2018). Prediction models for the incidence and progression of periodontitis: A systematic review. Journal of Clinical Periodontology, 45(12), 1408–1420.

Eickholz, P., Runschke, M., Dannewitz, B., Nickles, K., Petsos, H., Kronsteiner, D., & Pretzl, B. (2021). Long-term prognosis of teeth with class III furcation involvement. Journal of Clinical Periodontology, 48(12), 1528–1536.

Harrison, C. J., & Sidey-Gibbons, C. J. (2021). Machine learning in medicine: A practical introduction to natural language processing. BMC Medical Research Methodology, 21(1), 158.

Kantarjian, H., & Yu, P. P. (2015). Artificial intelligence, big data, and can- cer. JAMA Oncology, 1(5), 573–574.

Krois, J., Graetz, C., Holtfreter, B., Brinkmann, P., Kocher, T., & Schwendicke, F. (2019). Evaluating modeling and validation strategies for tooth loss. Journal of Dental Research, 98(10), 1088–1095.

Leow, N. M., Moreno, F., Marletta, D., Hussain, S. B., Buti, J., Almond, N., & Needleman, I. (2022). Recurrence and progression of periodontitis and methods of management in long-term care: A systematic review and meta-analysis. Journal of Clinical Periodontology, 49(Suppl 24), 291–313.

Mohammad-Rahimi, H., Motamedian, S. R., Pirayesh, Z., Haiat, A., Zahedrozegar, S., Mahmoudinia, E., Rohban, M. H., Krois, J., Lee, J. H., & Schwendicke, F. (2022). Deep learning in periodontology and oral implantology: A scoping review. Journal of Periodontal Research, 57, 942–951.

Ngiam, K. Y., & Khor, I. W. (2019). Big data and machine learning algo- rithms for health-care delivery. The Lancet Oncology, 20(5).

Phan, H. T. N. (2024). EARLY DETECTION OF ORAL DISEASES USING MACHINE LEARNING: A COMPARATIVE STUDY OF PREDICTIVE MODELS AND DIAGNOSTIC ACCURACY. International Journal of Medical Science and Public Health Research, 5(12), 107-118.

PHAN, H. T. N., & AKTER, A. (2024). HYBRID MACHINE LEARNING APPROACH FOR ORAL CANCER DIAGNOSIS AND CLASSIFICATION USING HISTOPATHOLOGICAL IMAGES. *Universal Publication Index e-Library*, 63-76.

Pretzl, B., Kaltschmitt, J., Kim, T. S., Reitmeir, P., & Eickholz, P. (2008). Tooth loss after active periodontal therapy. 2: Tooth-related factors. Journal of Clinical Periodontology, 35(2), 175–182.

Ravida, A., Troiano, G., Qazi, M., Saleh, M. H. A., Lo Russo, L., Greenwell, H., Giannobile, W. V., & Wang, H. L. (2020). Development of a nomogram for the prediction of periodontal tooth loss using the staging and grading system: A long-term cohort study. Journal of Clinical Periodontology, 47(11).

Saleh, M. H. A., Dukka, H., Troiano, G., Ravida, A., Galli, M., Qazi, M., Wang, H. L. (2021). External validation and comparison of the predictive performance of 10 different tooth-level prognostic systems. Jour- nal of Clinical Periodontology, 48(11), 1421–1429.

Saleh, M. H. A., Dukka, H., Troiano, G., Ravida, A., Qazi, M., Wang, H. L., & Greenwell, H. (2022). Long term comparison of the prognostic performance of PerioRisk, periodontal risk assessment, periodontal risk cal- culator, and staging and grading systems. Journal of Periodontology, 93(1), 57–68.

Saydzai, S., Buontempo, Z., Patel, P., Hasan, F., Sun, C., Akcali, A., Lin, G. H., Donos, N., & Nibali, L. (2022). Comparison of the efficacy of periodontal prognostic systems in predicting tooth loss. Journal of Clinical Periodontology, 49(8), 740–748.

Schwendicke, F., Schmietendorf, E., Plaumann, A., Salzer, S., Dorfer, C. E., & Graetz, C. (2018). Validation of multivariable models for predicting tooth loss in periodontitis patients. Journal of Clinical Periodontology, 45(6), 701–710.

Schwendicke, F., Stolpe, M., & Graetz, C. (2017). Cost comparison of prediction-based decision-making for periodontally affected molars. Journal of Clinical Periodontology, 44(11), 1145–1152.

Shi, S., Meng, Y., Li, W., Jiao, J., Meng, H., & Feng, X. (2020). A nomogram prediction for mandibular molar survival in Chinese patients with periodontitis: A 10-year retrospective cohort study. Journal of Clinical Periodontology, 47(9), 1121–1131.

Sidey-Gibbons, J. A. M., & Sidey-Gibbons, C. J. (2019). Machine learning in medicine: A practical introduction. BMC Medical Research Methodology, 19(1), 64.

Tusher, M. I., Phan, H. T. N., Akter, A., Mahin, M. R. H., & Ahmed, E. (2025). A Machine Learning Ensemble Approach for Early Detection of Oral Cancer: Integrating Clinical Data and Imaging Analysis in the Public Health. International Journal of Medical Science and Public Health Research, 6(04), 07-15.