Journal of Neonatal Surgery
*ISSN(Online): 2226-0439*
*Vol. 14, Issue 32s (2025)*
https://www.jneonatalsurg.com
OPEN ACCESS

# Automated Detection and Classification of Ponticulus Posticus from Digital Lateral Cephalograms- An Artifical Intelligence Based Retrospective Study

**Dr. Jebarani Jeevitha[1], Dr. Lokesh Kumar S[*2]**

[1]Department of Oral Medicine, Radiology, and Special Care Dentistry, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, 162, Poonamalle High Road, Chennai- 600077, India.

Email ID: jebaranijeevitha@gmail.com

[*2]Senior Lecturer, Department of Oral Medicine, Radiology, and Special Care Dentistry, Saveetha Dental College and Hospitals, SaveethaInstitute of Medical and Technical Sciences (SIMATS), Saveetha University, 162, Poonamallee High Road, Chennai- 600077, India.

[*]**Corresponding author:**

Dr. Lokesh Kumar S,

Email ID: lokeshkumars.sdc@saveetha.com

Cite this paper as: Dr. Jebarani Jeevitha, Dr. Lokesh Kumar S, (2025) Automated Detection and Classification of Ponticulus Posticus from Digital Lateral Cephalograms- An Artifical Intelligence Based Retrospective Study. *Journal of Neonatal Surgery*, 14 (32s), 7417-7424.

## ABSTRACT

**Context:** Ponticulus Posticus (PP) is a clinically significant anatomical variant of vertebrae that may compress the vertebral artery and contribute to cervicogenic headaches. Traditional radiographic assessment of PP is often subjective and inconsistent. This study explores the application of artificial intelligence (AI) for detection and classification of PP from digital lateral cephalograms.

**Aim:** To evaluate the accuracy of AI in detecting and classifying Ponticulus Posticus from digital lateral cephalograms.

**Settings and Design:** An artificial intelligence based retrospective study.

**Methods and materials:** A total of 1052 digital lateral cephalograms were selected, analyzed, and grouped as complete PP, partial PP, or absence of PP. Machine learning models in Orange® software, including Logistic Regression, Neural Network, and Naïve Bayes, were used for detection and classification.

**Results:** Logistic Regression achieved an accuracy of 98.5%, 98%, and 97.4% in the detection and classification of complete PP, partial PP, and absence of PP, respectively, outperforming Neural network and Naïve Bayes. Logistic Regression also demonstrated higher AUC, F1 score, precision, and recall. ROC curve analysis confirmed its superior classification ability across all PP categories.

**Conclusions:** AI, particularly Logistic Regression algorithm, is a reliable and promising tool for detecting and classifying Ponticulus Posticus in digital lateral cephalograms. Further validation using larger and more diverse datasets is recommended to enhance diagnostic precision.

***Keywords:*** *Artificial Intelligence; Ponticulus Posticus; Lateral Cephalogram; Machine Learning; Technology; Machine Learning*

## 1. INTRODUCTION

Ponticulus Posticus (PP), also known as the arcuate foramen, is a bony anatomical variation resulting from ossification of the posterior atlanto-occipital membrane over the groove for the vertebral artery on the atlas (C1) vertebra[1]. The presence of PP can alter the normal biomechanical function of the craniovertebral junction and has been implicated in neurovascular compression syndromes, including cervicogenic headaches, migraines, vertigo, and vertebrobasilar insufficiency.[2,3] In surgical and interventional settings—such as cervical spine instrumentation or atlanto-occipital fusion—failure to recognize a complete or partial PP may lead to inadvertent injury of the vertebral artery or adjacent neural structures.[4]

Epidemiological investigations report a wide variability in the prevalence of PP, ranging from 5.1% to 37.8% across different populations.[1,5] These differences are attributable to factors such as ethnicity, age group, and the imaging modality employed. Some studies observe a higher occurrence in females and in older age cohorts, suggesting a role for age-related calcification or mechanical loading in the progression of PP.[2,6] Morphologically, PP presents in two main forms: complete, when a continuous bony ring is evident, and partial, when ossification is incomplete, leaving gaps in the arch.

In routine dental and orthopedic practice, lateral cephalometric radiographs remain the first-line imaging modality for detecting PP due to their accessibility and low radiation dose[7]. However, manual evaluation is inherently subjective and prone to substantial intra- and inter-observer variability.[5] The presence of subtle or incomplete ossifications may be overlooked, particularly during high-volume reporting sessions. Moreover, radiologist fatigue, repetitive visual tasks, and cognitive overload can further compromise diagnostic reliability, increasing the risk of misclassification or oversight in clinical workflows.[7] These limitations underscore the need for more objective, reproducible diagnostic tools.

Artificial intelligence (AI) and machine learning (ML) have demonstrated considerable promise in medical image analysis, offering the ability to process large datasets, extract complex features, and deliver consistent classifications. Models such as Logistic Regression and Naive Bayes can be trained to recognize radiographic patterns of PP—complete, partial, or absent—potentially reducing observer bias and improving diagnostic throughput. Despite encouraging results in other areas of maxillofacial imaging, the diagnostic performance of AI algorithms for PP detection on lateral cephalograms has not been comprehensively validated. It remains unclear whether AI can match or exceed manual interpretation in terms of accuracy, sensitivity, and specificity.

Therefore, the present study was designed to evaluate the diagnostic accuracy of AI in detecting and classifying Ponticulus Posticus from digital lateral cephalograms. The aim of this research is to assess how accurately AI-based models can identify PP and distinguish between complete, partial, and absent forms. The objective is to compare the performance metrics—accuracy, precision, recall, F1 score, and AUC- of Logistic Regression and Naive Bayes algorithms against conventional manual radiographic evaluation. By investigating these comparative outcomes, the present study seeks to determine whether AI can enhance diagnostic consistency, reduce observer-dependent variability, and ultimately support clinical decision-making in the assessment of Ponticulus Posticus.

## 2. MATERIALS AND METHODS

This retrospective study was conducted over a six-month period (January–June 2024) at the Department of Oral Medicine and Radiology, at a private teaching instituition in Chennai. Ethical clearance was obtained from the Institutional Review Board prior to the study (IRB No. SDCH/2023/045). The required sample size was calculated using the formula $n = Z^2 p (1 - p)/d^2$, assuming an expected PP prevalence (p) of 0.162[8], precision (d) of 0.05, and 99% confidence (Z = 2.576), yielding a minimum n = 361; to account for possible exclusions. The digital lateral cephalograms of patients aged more than 10 years and less than 40 years, with no history of trauma or bone diseases and without any distortions were selected. However, the lateral cephalograms of patients below 10 or above 40 years of age, images with a history of cervical trauma or bone disorders, radiographs exhibiting distortion or poor visualization of the posterior arch (e.g., overlap by mastoid or occiput), and scans with technical artefacts were excluded.

A total of 1052 lateral cephalogram images made using a single OPG/ Cephalostat machine- Carestream 8100 running Carestream Dental Imaging Software (Carestream Dental LLC, 2022, USA) were selected following the inclusion and exclusion criteria. All radiographs were manually evaluated by a single calibrated experienced radiologist to minimize inter-observer variability and no more than 10 images per day were evaluated to avoid eye fatigue. Each image was assessed for the presence of complete PP (fully ossified bony ring), partial PP (incomplete ossification), or absence of PP.

Preprocessing of the selected images—including cropping to the atlas region and resizing to 256 × 256 pixels—was performed using Carestream Dental Imaging Software (Carestream Dental LLC, 2022, USA). Brightness and contrast normalization were applied to reduce lighting variability. The preprocessed images were then imported into Orange® Data Mining Software (University of Ljubljana, 2023, Slovenia)[9] for feature extraction, where texture, shape, and intensity features were converted into numerical vectors to train and test the models.

Three machine learning classifiers—Logistic Regression, Neural network, and Naïve Bayes algorithms were trained on 75% of the dataset and tested on the remaining 25%. Model performance was evaluated using classification accuracy (the proportion of correct predictions), F1-score (the harmonic mean of precision and recall), precision (the proportion of true positives among all positive predictions), recall (sensitivity), Matthew's correlation coefficient (MCC), where it correlates the predicted & the actual binary outcomes, and area under the receiver operating characteristic curve (AUC), where higher AUC indicates better discrimination between classes.

## 3. RESULTS

**Distribution of Ponticulus Posticus:**

A total of 1052 lateral cephalograms were analyzed, comprising 407 males and 645 females. The overall distribution by gender and PP category is as follows:

- Complete PP: 14 males (3.4 %), 24 females (3.7 %)
- Partial PP: 44 males (10.8 %), 89 females (13.8 %)
- Absence of PP: 349 males (85.7 %), 532 females (82.5 %)

Thus, the combined prevalence of PP (complete + partial) was 14.3 % in males (58/407) and 17.5 % in females (113/645).

**Confusion matrix analysis:**

The confusion matrix of the three algorithms under the study are depicted in table 1. The true positive rate (sensitivity), true negative rate (specificity), false positive rate and false negative rate values for the three algorithms are as follows:

| **Algorithm** | **Confusion matrix** | | | | | |
|---|---|---|---|---|---|---|
| **Logistic Regression** | | **Predicted** | | | | |
| | | | **Complete PP** | **Absence of PP** | **Partial PP** | **Σ** |
| | **Actual** | **Complete PP** | 26 | 9 | 3 | **38** |
| | | **Absence of PP** | 2 | 874 | 5 | **881** |
| | | **Partial PP** | 2 | 11 | 120 | **133** |
| | | **Σ** | **30** | **894** | **128** | **1052** |
| **Neural Network** | | **Predicted** | | | | |
| | | | **Complete PP** | **Absence of PP** | **Partial PP** | **Σ** |
| | **Actual** | **Complete PP** | 23 | 11 | 4 | **38** |
| | | **Absence of PP** | 3 | 872 | 6 | **881** |
| | | **Partial PP** | 2 | 11 | 120 | **133** |
| | | **Σ** | **28** | **894** | **130** | **1052** |
| **Naïve Bayes** | | **Predicted** | | | | |
| | | | **Complete PP** | **Absence of PP** | **Partial PP** | **Σ** |
| | **Actual** | **Complete PP** | 33 | 3 | 2 | **38** |
| | | **Absence of PP** | 283 | 584 | 14 | **881** |
| | | **Partial PP** | 19 | 9 | 105 | **133** |
| | | **Σ** | **335** | **596** | **121** | **1052** |

**Table 1: Confusion matrix for the detection and classification of Ponticulus Posticus by Logistic regression, Neural network, and Naïve Bayes algorithms**

**Table 2: Test and score model performance results of AI algorithms for complete, partial, and absence of Ponticulus Porticus**

| Type of PP | Model | AUC | CA | F1 | Precision | Recall | MCC |
|---|---|---|---|---|---|---|---|
| **Complete PP** | **Logistic Regression** | 0.949 | 0.985 | 0.765 | 0.867 | 0.684 | 0.763 |
| | **Neural Network** | 0.918 | 0.981 | 0.697 | 0.821 | 0.605 | 0.696 |
| | **Naïve Bayes** | 0.830 | 0.830 | 0.177 | 0.099 | 0.868 | 0.229 |
| **Partial PP** | **Logistic Regression** | 0.977 | 0.980 | 0.920 | 0.938 | 0.902 | 0.908 |
| | **Neural Network** | 0.964 | 0.978 | 0.913 | 0.923 | 0.902 | 0.900 |
| | **Naïve Bayes** | 0.927 | 0.958 | 0.827 | 0.868 | 0.789 | 0.804 |
| **Absence of PP** | **Logistic Regression** | 0.972 | 0.974 | 0.985 | 0.978 | 0.992 | 0.904 |
| | **Neural Network** | 0.961 | 0.971 | 0.983 | 0.975 | 0.990 | 0.889 |
| | **Naïve Bayes** | 0.942 | 0.706 | 0.791 | 0.980 | 0.663 | 0.441 |

**PP- Ponticulus Posticus; AUC- Area under the curve; CA- Classification Accuracy; MCC- Matthew's Correlation Coefficient.**

Logistic regression:

- TP: 26 (complete) + 120 (partial) = 146 of 171 PP cases → Sensitivity= 85.4 %
- TN: 874 of 881 non-PP (absence) cases → Specificity= 99.2 %
- FP: 7 of 881 non-PP (absence) cases → false-positive rate= 0.8 %
- FN: 25 of 171 PP cases missed → false-negative rate= 14.6 %

Neural Network:

- TP: 23 (complete) + 120 (partial) = 143 of 171 PP cases → Sensitivity= 83.6 %
- TN: 872 of 881 non-PP cases (absence) → Specificity= 99.0 %
- FP: 9 of 881 non-PP cases (absence) → false-positive rate (FPR)= 1.0 %
- FN: 22 of 171 PP cases missed → false-negative rate (FNR)= 16.4 %

Naïve Bayes:

- TP: 33 (complete) + 105 (partial) = 138 of 171 actual PP cases → Sensitivity= 92 %
- TN: 584 of 881 non-PP cases (absence) → Specificity= 66.3 %
- FP: 297 of 881 non-PP cases (absence) → false-positive rate (FPR)= 33.7 %
- FN: 12 of 171 PP cases missed → false-negative rate (FNR)= 8 %

Hence, Naïve Bayes had the highest sensitivity among the three algorithms, whereas showed a low specificity. However, Logistic regression and Neural network algorithms displayed excellent specificity and an acceptable good sensitivity. Logistic regression algorithm had a good balance between the two parameters among the three algorithms.

**Performance of the machine learning models:**

Three machine learning models—Logistic Regression, Neural Network, and Naive Bayes—were evaluated for their ability to classify lateral cephalograms into the three PP categories. The performance of each model was assessed using key metrics: Area Under the Curve (AUC), Classification Accuracy (CA), F1 Score, Precision, Recall, and Matthews Correlation Coefficient (MCC).

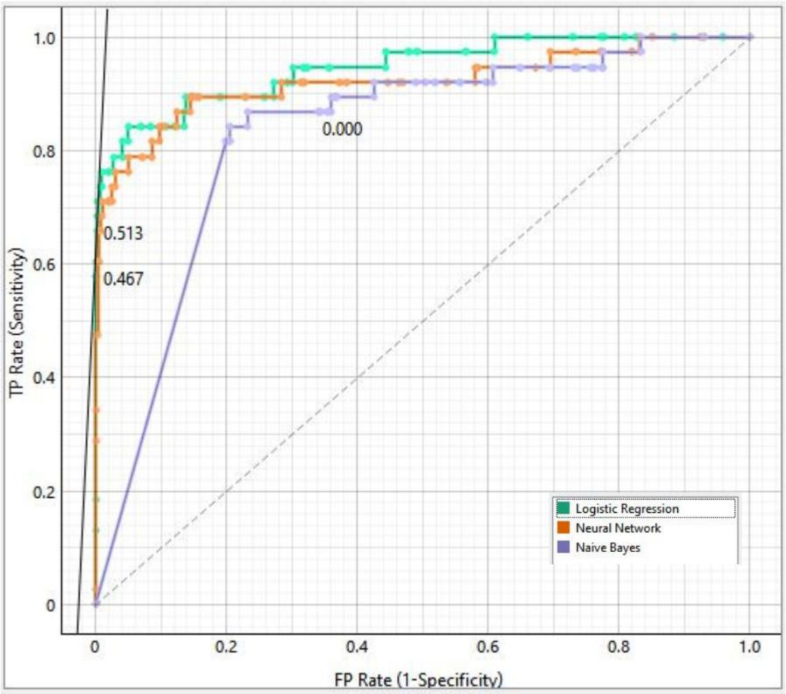
**Complete PP:** The Logistic Regression model achieved an accuracy of 98.5%, an AUC of 0.949, and an F1 score of 0.765—substantially higher than the others. Neural Network performance was modest here with an F1 score of 0.697 and Recall of 60.5%, while Naive Bayes showed the weakest results, with an F1 score of just 0.177, Precision of 0.099, and Recall of only 8.6%. Across all categories, Logistic Regression consistently achieved the highest AUC, accuracy, F1 score, precision, and recall.

**Partial PP:** Logistic Regression again led the performance charts, achieving AUC of 0.977, accuracy of 98.0%, and F1 score of 0.920. The Neural Network model showed competitive performance with an accuracy of 97.8%, F1 score of 0.913, and AUC of 0.964. Naive Bayes lagged behind with an accuracy of 95.8% and a lower F1 score of 0.827, confirming its limitations in distinguishing Partial PP cases.

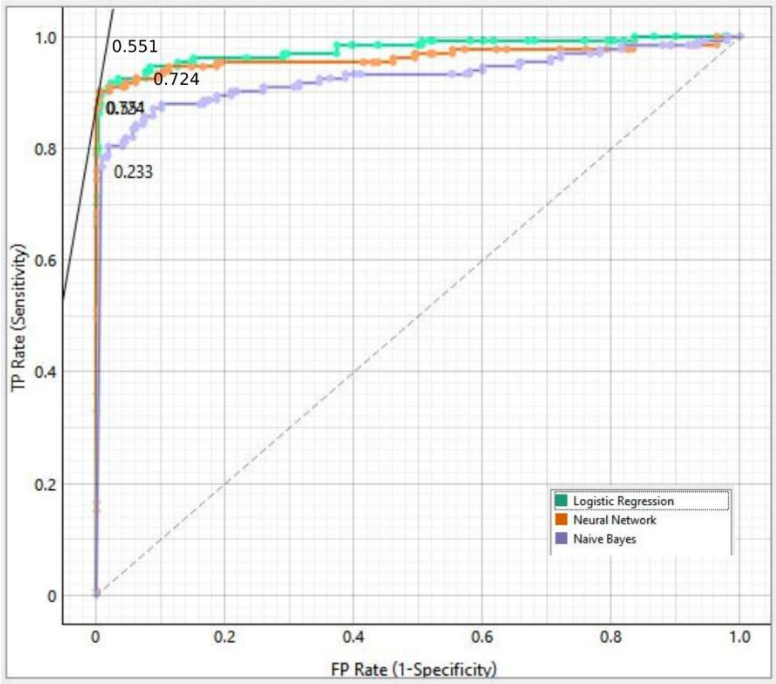
**Absence of PP:** Logistic Regression outperformed all other models with an AUC of 0.972, accuracy of 97.4%, and F1 score of 0.985, indicating high reliability and consistency. Neural Network followed closely with an AUC of 0.961, accuracy of 97.1%, and F1 score of 0.983. Naive Bayes, however, performed significantly worse, with a lower accuracy of 70.6%, F1 score of 0.791, and Recall of only 66.3%, reflecting high misclassification rates and poor specificity.

**Receiver Operating Characteristic (ROC) Curve Analysis:**

ROC curve analysis provided further insight into the discriminatory power of each model for the three classification tasks.

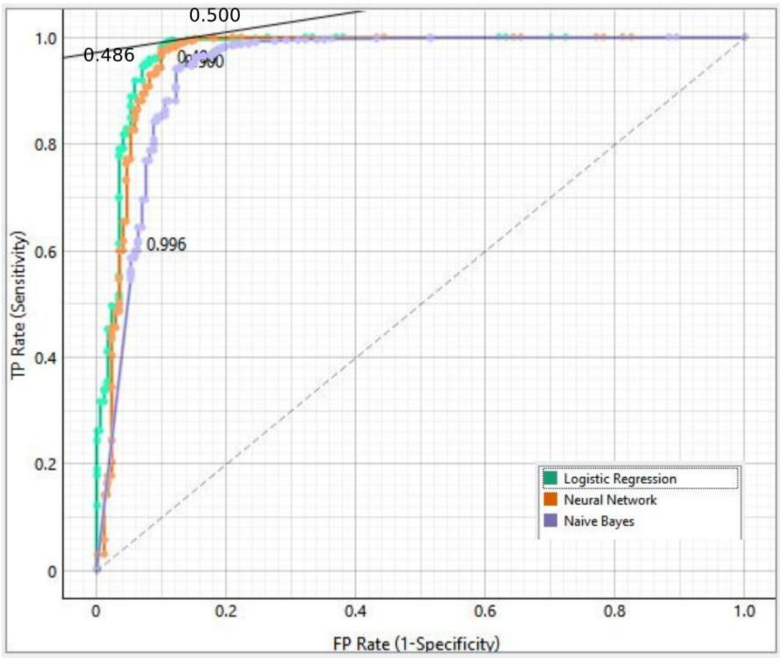



**Complete PP (Figure 1): Despite the inherent complexity of detecting complete PP due to its low prevalence, Logistic Regression continued to outperform the other models. Neural Network showed reasonable ROC characteristics, while Naive Bayes had the most flattened curve, affirming its weak classification ability in this category.**




**Partial PP (Figure 2): Logistic Regression displayed the most favorable ROC characteristics, confirming its precision in detecting partial ossifications. The Neural Network model showed slightly lower sensitivity, while Naive Bayes had visibly inferior classification performance, especially at low false positive rates.**

The Logistic Regression curve lies closest to the top-left corner—corresponding to its superior AUC values (0.949, 0.977, and 0.972) for complete, partial, and absence of PP, respectively). In contrast, Neural Network AUCs range from 0.918 to 0.964, and Naive Bayes from 0.830 to 0.942. All models performed well above the diagonal (random classification), but Logistic Regression demonstrates the greatest discrimination across thresholds.




**Absence of PP (Figure 3): The ROC curve for Logistic Regression showed the steepest ascent and highest proximity to the top-left corner, indicating excellent sensitivity and specificity. Both Neural Network and Naive Bayes lagged slightly behind, with the latter displaying lower curve elevation and AUC.**

## 4. DISCUSSION

The current study assessed the diagnostic accuracy of three machine learning algorithms—Logistic Regression, Neural Network, and Naive Bayes—for detecting and classifying Ponticulus Posticus (PP) on digital lateral cephalograms, comparing each model's output against manual radiographic evaluation. Unlike prior work that assumed AI capability, we explicitly evaluated model performance in categorizing PP as complete, partial, or absent.

In our South Indian cohort aged 10–40 years, the combined complete and partial PP prevalence was 16.6 % (171/1,052), as compared to the global prevalence range of 5–45.5 %.[10] A meta-analysis and review of literature by Elliott RE and Tanweer O has also reported the overall prevalence to be 16.7% among 22000 patients.[8] However, the prevalence in the present study may reflect ethnic, developmental, and imaging-modality differences, as lateral cephalograms can underestimate bony bridges compared to cone-beam computed tomography.

Etiologically, PP formation is multifactorial. Congenital ossification of the atlanto-occipital membrane, age-related calcification, and mechanical loading on the cervical spine each contribute to bony arch development.[7,11] Genetic predispositions likely establish the initial ossification pattern, whereas repetitive strain—common in certain occupations—may accelerate progression.[11] By limiting inclusion to patients ≤ 40 years, we minimized age-related bias yet still observed PP in adolescents, underscoring its developmental origins.[4]

Gender differences were evident: PP occurred in 14.3 % of males and 17.5 % of females, indicating a modest female predilection. This concurs with earlier reports which noted similar gender biases linked to hormonal and anatomic variations.[6,12] Given the higher incidence of migraine and cervical pain among women, recognizing this disparity is clinically significant.

Several studies have applied the same algorithms to different craniofacial parameters. One study used Logistic Regression to detect sella turcica bridging with 95 % accuracy[13], while another applied Neural Networks to classify temporomandibular joint disorders, achieving an AUC of 0.96.[14] Naive Bayes has been used for osteoporosis screening on panoramic radiographs, with 85 % sensitivity.[15] These investigations, alongside the previous styloid process study[16], establish a precedent for AI in dental radiology and justify our algorithmic choices.

In the present study, Logistic Regression achieved the highest overall accuracy, F1 score and AUC outperforming Naive Bayes and Neural Network. It also had a good balance between sensitivity and specificity. Neural network showed a moderate performance. Notably, Naive Bayes, even though having a high sensitivity than other algorithms, tended to misclassify absence of PP as complete and partial PP, resulting in a high false positive rate and low specificity. Logistic Regression's transparent decision boundaries facilitate clinical interpretability—an essential feature for adoption—while maintaining robustness to class imbalance.[5] Similar findings were noted in the study by Jeevitha SJ et al where logistic regression depicted higher performance than the other two algorithms namely Neural Network and Naïve Bayes.[16]

The precision-recall trade-off further favored Logistic Regression for high-sensitivity applications, whereas Naive Bayes may require feature enhancement or data-balancing strategies. Future work should explore ensemble approaches, multi-modal imaging, and integration of clinical symptom data to further refine AI-assisted PP detection.

## 5. LIMITATIONS

The dataset used in this study was sourced from a single institution and may not fully represent broader population diversity. The performance of the AI models was not validated on external datasets, which may affect generalizability.

## 6. FUTURE PROSPECTS

Future studies should explore the integration of clinical symptoms with radiographic data to enhance diagnostic precision. Studies involving multicenter, multi-ethnic datasets and comparisons with three-dimensional imaging modalities like cone beam computed tomography (CBCT) could improve the robustness and applicability of the findings.

## 7. CONCLUSION

Logistic Regression consistently achieved the highest performance across all metrics, confirming its diagnostic strength. Neural Network demonstrated moderate capability, whereas Naive Bayes performed poorly, especially in distinguishing normal scans and detecting complete PP. These findings underscore the importance of selecting suitably balanced and interpretable models for clinical applications where precision is critical. Also, the incorporation of Artificial Intelligence and Machine Learning approaches in the detection of Ponticulus Posticus holds a promising value.

## REFERENCES

[1] Cho YJ. Radiological analysis of ponticulus posticus in Koreans. Yonsei Med J 2009; 50(1):45–9.

[2] Govindraju P, Kumar TSM. Prevalence of ponticulus posticus of the first cervical vertebra: a digital radiographic study. J Indian Acad Oral Med Radiol 2017; 29(2):95–9.

[3] Suazo GI, Schilling J, Schilling A. Ponticulus posticus on the posterior arch of atlas: prevalence analysis in asymptomatic patients. Int J Morphol 2010;28(1):317–22.

[4] Chitroda PK, Katti G, Sreedevi E, Mandlik J, Pawar R, Shetty S. Ponticulus posticus on the posterior arch of atlas: prevalence analysis in symptomatic and asymptomatic patients of Gulbarga population. J Clin Diagn Res 2013; 7(12):3044–7.

[5] Diallo R, Edalo C, Awe O. Machine learning evaluation of imbalanced health data: a comparative analysis of balanced accuracy, MCC, and F1 score. In: Arai K, editor. Proceedings of the Future of Information and Communication Conference (FICC) 2024. Cham: Springer; 2024. p. 283–312.

[6] Macri M, Rendina F, Marini M, Papi P, Pompa G. Prevalence of ponticulus posticus and migraine in 220 orthodontic patients: a cross-sectional study. Biology (Basel) 2023; 12(3):471.

[7] Tassoker M, Kok H, Ozcan S. Investigation of the relationship between sella turcica bridge and ponticulus posticus: a lateral cephalometric study. Int J Morphol 2017; 35(2):337–44.

[8] Elliott RE, Tanweer O. The prevalence of the ponticulus posticus (arcuate foramen) and its importance in the Goel-Harms procedure: meta-analysis and review of the literature. World neurosurg 2014; 82(1-2):e335-43.

[9] Data mining fruitful And fun. (1996). Accessed: July 09, 2025: https://orangedatamining.com/.

[10] Xu X, Zhu Y, Ding X, Yin M, Mo W, Ma J. Research Progress of Ponticulus Posticus: A Narrative Literature Review. Front Surg 2022; 9:834551. doi: 10.3389/fsurg.2022.834551.

[11] Shahidi S, Hasani M, Khozaei M. Evaluating the relation between the elongated styloid process and the ponticulus posticus using cone-beam computed tomography. Folia Morphologica 2022; 81(1):196-202.

[12] Sharma V, Chaudhary D, Mitra R. Prevalence of ponticulus posticus in Indian orthodontic patients. Dentomaxillofac Radiol 2010; 39(5):277-83.

[13] Singh R. Logistic regression-based detection of sella turcica bridging on cephalograms. Dentomaxillofac Radiol 2022; 51(2):20210235.

[14] Singh P, Gupta V, Sharma S. Neural network classification of temporomandibular joint disorders on MRI images. J Oral Rehabil 2021; 48(7):826–33.

[15] Kumar N, Verma S, Rai A. Naive Bayes algorithm for screening osteoporosis on panoramic radiographs. Oral Radiol 2020; 36(4):365–72.

[16] Jeevitha S J, Kumar L, Yadalam PK, kumar Yadalam P. Artificial intelligence: a reliable tool to detect the elongation of the styloid process. Cureus 2023; 15(11):e94541.