



www.bioinformation.net
Volume 21(6)



Research Article

Received June 01, 2025; Revised June 30, 2025; Accepted June 30, 2025, Published June 30, 2025

DOI: 10.6026/973206300211555

SJIF 2025 (Scientific Journal Impact Factor for 2025) = 8.478

2022 Impact Factor (2023 Clarivate Inc. release) is 1.9

Declaration on Publication Ethics:

The author's state that they adhere with COPE guidelines on publishing ethics as described elsewhere at <https://publicationethics.org/>. The authors also undertake that they are not associated with any other third party (governmental or non-governmental agencies) linking with any form of unethical issues connecting to this publication. The authors also declare that they are not withholding any information that is misleading to the publisher in regard to this article.

Declaration on official E-mail:

The corresponding author declares that lifetime official e-mail from their institution is not available for all authors

License statement:

This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License

Comments from readers:

Articles published in BIOINFORMATION are open for relevant post publication comments and criticisms, which will be published immediately linking to the original article without open access charges. Comments should be concise, coherent and critical in less than 1000 words.

Disclaimer:

Bioinformation provides a platform for scholarly communication of data and information to create knowledge in the Biological/Biomedical domain after adequate peer/editorial reviews and editing entertaining revisions where required. The views and opinions expressed are those of the author(s) and do not reflect the views or opinions of Bioinformation and (or) its publisher Biomedical Informatics. Biomedical Informatics remains neutral and allows authors to specify their address and affiliation details including territory where required.

Edited by Hiroj Bagde MDS, (PhD), PGDCR, PGDHHM, PGDL, PGDM

E-mail: hirojbagde8@gmail.com; Phone: +91 9766105900

Citation: Samir *et al.* Bioinformation 21(6): 1555-1558 (2025)

Behavioral outcome prediction among children using machine learning

P.V. Samir^{1,*}, G Aruna Kumari², Nandini Biradar³, Kodali Srija², Debasmita Das⁴ & Sukabhogi Anusha⁵

¹Department of Pedodontics & Preventive Dentistry, Kalinga Institute of Dental Sciences, KIIT Deemed to be University, Bhubaneswar-751006, India; ²Department of Pedodontics & Preventive Dentistry, Malla Reddy Dental College for Women-Mallareddy Viswavidyapeeth, Hyderabad, India; ³Department of Conservative dentistry and Endodontics, Private Practitioner, Biradar Dental care, Bidar, Karnataka, India; ⁴Department of Pedodontics and Preventive Dentistry, Hi-Tech Dental College & Hospital, Bhubaneswar, India; ⁵Department of Pedodontics and Preventive Dentistry, Meghana Institute of Dental Sciences, Nizamabad- 503001, Telangana, India; *Corresponding author

Affiliation URL:

<https://kids.kiit.ac.in/>
<https://mrdcw.edu.in/>
<https://hi-techdentalbbsr.org/>
<http://meghnadentalcollege.ac.in/>

Author contacts:

PV Samir - E-mail: pv.samir@kids.ac.in;
G Aruna Kumari - E-mail: aruna2thdoc@gmail.com;
Nandini Biradar - E-mail: dr.nandini.salger@gmail.com;
Kodali Srija - E-mail: srijakodali5@gmail.com
Debasmita Das - E-mail: dasdebasmita08@gmail.com;
Sukabhogi Anusha - E-mail: anushapatil.sukhabogi@gmail.com;

Abstract:

Behavioural management in paediatric dentistry is essential for treatment success, yet predicting a child's behavior remains a challenge. This study used machine learning models on data from 120 children aged 4–10 years, incorporating clinical and historical variables such as age, dental history and parental anxiety. Among the models tested, Random Forest achieved the highest accuracy (87.5%) in predicting behavior based on the Frankl scale. Key predictors of negative behavior included younger age, high parental anxiety and prior negative dental experiences. These findings highlight the potential of machine learning to support behavior guidance planning and improve clinical outcomes.

Keywords: Pediatric dentistry, machine learning, behavioral prediction, dental anxiety, random forest, frankl scale

Background:

The accurate prediction of paediatric cognitive and behavioural outcomes remains a significant challenge in clinical and developmental neuroscience. Traditional diagnostic pathways often fail to identify high-risk children early enough for effective intervention. Recent advancements in machine learning (ML) and big data analytics have shown promising potential in addressing this gap through predictive modelling across a range of paediatric conditions [1, 2]. Several studies have demonstrated the utility of ML in predicting behavioural and neurodevelopmental outcomes using diverse data inputs including clinical, behavioural, physiological and socio-demographic variables. For instance, ML algorithms using electronic health records have effectively predicted clinical outcomes during emergency triage in children, outperforming conventional tools in risk stratification and care planning [2]. Similarly, ML approaches have successfully identified behavioural indicators for ADHD [3], disruptive behaviour via physiological smart-watch data [4] and future onset of bipolar disorder based on childhood behavioural features [5, 13]. In pediatric traumatic brain injury, ML models have exceeded the accuracy of CT-based systems in predicting long-term neurological outcomes [6, 12]. Autism spectrum disorder (ASD) has also been a focal point, with studies showing that ML can reliably forecast treatment responses and behavioural challenges based on sensory processing and executive function profiles [7–9, 15]. In psychiatric contexts, ML has predicted suicide risk [10], post-traumatic stress [14] and outcomes from internet-based CBT in obsessive-compulsive disorder [11].

Building on these applications, researchers have explored the use of ML to predict cognitive development trajectories by

incorporating neurocognitive assessments, genetic data and environmental exposures. Behera *et al.* [1] utilized an interdisciplinary dataset comprising socioeconomic factors, parental history and cognitive testing outcomes, demonstrating that predictive models could identify children at risk for poor cognitive performance well before clinical symptoms manifest. The integration of such heterogeneous data types improves the robustness of predictive models, ensuring higher sensitivity and specificity in developmental risk profiling. Moreover, ML models have also shown superiority over traditional statistical approaches in intervention outcome prediction. In a quasi-experimental setting, Sun *et al.* [8] demonstrated that ML techniques, particularly support vector machines, predicted which children with autism would benefit most from structured motor skill interventions such as mini-basketball training. These findings suggest that ML can not only support diagnostic predictions but also enhance personalized treatment planning by identifying likely responders to specific therapies. As data availability and computational power continue to increase, ML holds transformative potential to shift pediatric care from reactive to proactive by offering early, data-driven insights into neurodevelopmental and behavioural disorders.

Materials and Methods:**Study population:**

The study included 120 children aged between 4 and 10 years who reported for routine dental treatment. Children with special healthcare needs or those on medications affecting behavior or cognition were excluded. Sampling was done using a non-probability convenience sampling method.

Data collection:

Data were collected through a structured form consisting of the following variables:

- [1] **Demographic details:** age, gender
- [2] **Dental history:** number of previous dental visits, experience during past dental treatment (positive/negative)
- [3] **Parental factors:** anxiety level measured using a Modified Corah’s Dental Anxiety Scale
- [4] **Clinical indicators:** caries index (DMFT/deft), oral hygiene status (Simplified Oral Hygiene Index) and subjective pain perception using a Faces Pain Scale

Behavioral assessment:

Child behavior during the dental procedure was assessed using the Frankl Behaviour Rating Scale, which classifies behavior into four categories: definitely negative, negative, positive and definitely positive. This assessment was conducted by a calibrated examiner who was blinded to the machine learning process.

Machine learning model development:

Data were pre-processed by handling missing values and converting categorical variables into numerical formats using one-hot encoding. The dataset was split into a training set (80%) and a testing set (20%). Three supervised machine learning algorithms were implemented:

- [1] Decision tree
- [2] Random forest
- [3] Support vector machine (SVM)

These models were developed using Python and the Scikit-learn library. The models were trained using the training dataset and evaluated on the testing set. The outcome variable was the child’s behavioral response (binary classification: positive vs. negative behavior).

Model evaluation:

Performance of the models was evaluated using standard classification metrics:

- [1] **Accuracy** = (True Positives + True Negatives) / Total Predictions
- [2] **Precision** = True Positives / (True Positives + False Positives)
- [3] **Recall (Sensitivity)** = True Positives / (True Positives + False Negatives)
- [4] **F1-score** = 2 × (Precision × Recall) / (Precision + Recall)

The model with the highest overall performance across all metrics was identified as the most effective for predicting behavioral outcomes.

Results:

Baseline characteristics:

A total of 120 children were included in the study, comprising 68 males (56.7%) and 52 females (43.3%). The mean age was 6.7 ± 1.9 years. A majority of participants (62.5%) had at least one prior dental visit and 41.6% reported a negative past dental experience. High parental anxiety scores were observed in 39.2% of cases. The distribution of clinical variables and behavioral outcomes is presented in **Table 1**. As shown in **Table 1**, 39.2% of children exhibited negative behavior during treatment. A higher proportion of negative behavior was observed among younger children, those with high parental anxiety and those reporting previous negative dental experiences.

Machine learning model performance:

Among the three models evaluated, the Random Forest classifier demonstrated the highest predictive accuracy. Performance metrics for all models are detailed in **Table 2**. As seen in **Table 2**, the Random Forest model achieved the highest accuracy (87.5%) and F1-score (85.1%), followed by the SVM model with an accuracy of 81.3%. The Decision Tree model showed moderate performance across all metrics. Key variables contributing to accurate predictions in the Random Forest model included age, parental anxiety and prior dental experience.

Table 1: Baseline demographic and clinical characteristics of the participants (n = 120)

Variable	Categories	Frequency (%)
Gender	Male / Female	68 (56.7%) / 52 (43.3%)
Age Group	4–6 / 7–10 years	64 (53.3%) / 56 (46.7%)
Previous Dental Visit	Yes / No	75 (62.5%) / 45 (37.5%)
Past Dental Experience	Positive / Negative	70 (58.4%) / 50 (41.6%)
Parental Anxiety (High)	Yes / No	47 (39.2%) / 73 (60.8%)
Caries Status (DMFT/deft >3)	Yes / No	79 (65.8%) / 41 (34.2%)
Oral Hygiene (Fair/Poor)	Fair / Poor	82 (68.3%) / 38 (31.7%)
Pain Perception (Moderate-Severe)	Yes / No	66 (55%) / 54 (45%)
Frankl Behavior (Positive)	Yes / No	73 (60.8%) / 47 (39.2%)

Table 2: Performance metrics of machine learning models for behavioral prediction

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Decision Tree	76.4	74.1	72.9	73.5
Random Forest	87.5	84.2	86.1	85.1
Support Vector Machine (SVM)	81.3	78.5	80.2	79.3

Discussion:

The results of this study reinforce the growing evidence that machine learning (ML) tools, particularly ensemble methods like Random Forests, can significantly enhance the prediction of behavioural outcomes in paediatric dental settings. The observed superiority of the Random Forest model across accuracy, precision, recall and F1-score echoes prior findings from broader pediatric research domains, where similar methods have consistently outperformed simpler algorithms such as decision trees or logistic regression [1, 2 and 13]. In our study, key variables influencing behavioural prediction—age, parental dental anxiety and prior dental experience—align closely with earlier investigations into behavioural and cognitive predictors across various pediatric conditions. For instance, younger age has been linked to behavioural deregulation and reduced procedural cooperation, a trend echoed in ADHD prediction using behavioural activity data by Maniruzzaman *et al.* [3] and in autism-related disruptive behavior prediction using physiological data by Romanowicz *et al.* [4]. The strong predictive power of parental anxiety is also well-established, with studies confirming its direct impact on child distress levels in clinical settings [6]. Importantly, our findings resonate with those of Behera *et al.* [1], who emphasized that incorporating multidimensional inputs such as familial, behavioural and clinical features leads to highly accurate cognitive outcome predictions in children. Although their study focused more broadly on cognitive development, the integration of socio-environmental and clinical variables mirrors the structured feature set used in the present research. Similarly, the wearable-derived biomarkers used by Romanowicz *et al.* [4] in predicting disruptive behavior indicate a shift toward real-time and multimodal behavioural prediction models, suggesting future directions for expanding behavioural data collection in dental clinics. This study's focus on using ML within a dental context distinguishes it from many prior works that emphasized neurological or psychiatric outcomes [5, 6, 10 and 11]. Nevertheless, the observed pattern—that ML models can handle complex variable interactions more effectively than conventional tools—is consistent with findings in pediatric traumatic brain injury prognosis [6, 12] and suicide risk stratification [10]. Our Random Forest model's robustness further reflects the value of using ensemble learning in healthcare, particularly where outcomes are influenced by interacting clinical and psychosocial variables. Unlike many previous works that used outcomes from long-term follow-up or neuroimaging, this study employed the Frankl Behaviour Rating Scale—a validated yet practical tool for chairside behaviour classification [14]. While the binary classification (positive vs. negative) simplifies modelling, it may not capture the full behavioural nuance observed during dental procedures. Studies like that by Lenhard *et al.* [11] on OCD treatment outcomes demonstrate the utility of continuous outcome modelling, which may benefit future pediatric dental ML applications as well. Furthermore, predictive modelling in ASD therapy outcomes, such as Sun *et al.*'s work with motor skill

interventions [8], reveals the promise of ML in intervention stratification. Our findings similarly support the use of behavior prediction for individualized treatment planning in dental care—where interventions like tell-show-do, parental counselling, or pharmacologic behavior control can be proactively applied based on predicted cooperation levels. The relatively high performance of all three tested ML models also emphasizes the importance of using cross-validated, well-tuned models even in modestly sized datasets. However, like previous studies limited by sample diversity or follow-up length [9, 12 and 15], our study's single-centre design and limited inclusion of psychosocial and environmental features (e.g., parenting style, education level, or temperament) highlight areas for expansion. Longitudinal data and real-time digital behavior tracking, as explored in Saxe *et al.*'s PTSD prediction work [14], may provide more robust and dynamic predictive frameworks for future dental settings.

Conclusion:

Machine learning models, particularly Random Forests, show strong potential in predicting pediatric dental behavior using routine clinical and historical data. Early identification of children at risk for negative behavior can enhance treatment planning and improve patient cooperation, making AI-driven tools valuable assets in pediatric dental practice.

References:

- [1] Behera NR *et al.* In: 2024 4th International Conference on Innovative Practices in Technology and Management (ICIPTM); 2024; Noida, India. p6.
- [2] Goto T *et al.* *JAMA Netw Open.* 2019 **2**:e186937. [PMID: 30646206]
- [3] Maniruzzaman M *et al.* *Appl Sci.* 2022 **12**:2737. [DOI: 10.3390/app12052737].
- [4] Romanowicz M *et al.* *J Child Adolesc Psychopharmacol.* 2023 **33**:387 [PMID: 37966360]
- [5] Uchida M *et al.* *J Psychiatr Res.* 2022 **156**:261. [PMID: 36274531].
- [6] Hale AT *et al.* *Neurosurg Focus.* 2018 **45**:E2. [PMID: 30453455].
- [7] Alateyat H *et al.* *Front Mol Neurosci.* 2022 **15**:889641. [PMID: 35615066]
- [8] Sun Z *et al.* *Int J Ment Health Promot.* 2022 **24**:143. [DOI:10.32604/ijmhp.2022.020075]
- [9] Usta MB *et al.* *Psychiatry Clin Psychopharmacol.* 2019 **29**:320. [DOI: 10.1080/24750573.2018.1545334].
- [10] Su C *et al.* *Transl Psychiatry.* 2020 **10**:413. [PMID: 33243979]
- [11] Lenhard F *et al.* *Int J Methods Psychiatr Res.* 2018 **27**:e1576. [PMID: 28752937].
- [12] Tunthanathip T *et al.* *Chin J Traumatol.* 2021 **24**:350. [PMID: 34284922].
- [13] Uchida M *et al.* *J Psychiatr Res.* 2022 **156**:261. [PMID: 36274531].
- [14] Saxe GN *et al.* *BMC Psychiatry.* 2017 **17**:223. [DOI: 10.1186/s12888-017-1384-1].
- [15] Linstead E *et al.* In: 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA); 2015; Miami, FL, USA. p. 418. [DOI: 10.1109/ICMLA.2015.214].