Machine Learning–Based Prediction of RAIR Abnormality in Pediatric Patients Using High-Resolution Anorectal Manometry

Author: Chaitali Chaudhari
Institution: SSM Health Cardinal Glennon Children's Hospital

Abstract

Pediatric gastrointestinal (GI) dysmotility disorders are common and often require high-resolution anorectal manometry (ARM) for physiologic assessment. This study evaluates machine-learning (ML) approaches to predict rectoanal inhibitory reflex (RAIR) abnormality using ARM-derived physiologic features. De-identified ARM data curated by Dhiren Patel at SSM Health Cardinal Glennon Children's Hospital were analyzed. Predictors included ARM pressure and relaxation metrics (resting pressures, IAS percent relaxation at multiple thresholds, relaxation durations, squeeze and push maneuvers, and rectal sensation). The outcome was binary RAIR status (0 = present, 1 = absent/abnormal). We evaluated LASSO logistic regression, SVM, decision tree, Random Forest, Gradient Boosting, and an ensemble model with hyperparameter tuning via 5-fold cross-validation. Ensemble and Random Forest models achieved the best performance; IAS relaxation metrics (especially at 40%) and resting pressure were the most informative predictors. These findings demonstrate the feasibility of ML-assisted ARM interpretation and identify physiologic markers that may support clinical decision-making.

Introduction

Gastrointestinal (GI) complaints account for a substantial portion of pediatric clinical care and specialist referrals. Functional anorectal disorders, including functional constipation, involve abnormalities in anorectal sensation, coordination, or reflexes without structural disease and affect a significant proportion of children. High-resolution anorectal manometry (ARM) provides detailed measurements of anorectal physiology—including resting anal pressure, internal anal sphincter (IAS) relaxation, voluntary squeeze, simulated defecation, rectal sensation, and the rectoanal inhibitory reflex (RAIR)—but its predictive value for identifying clinically meaningful abnormalities in the pediatric population is not well-established.

Methods

Data Source and Attribution De-identified ARM studies and linked clinical fields were curated and provided by Dhiren Patel at SSM Health Cardinal Glennon Children's Hospital.

Predictors (ARM features) Predictors included all ARM-derived variables from ARMRestPresMean through Rectalsensation (inclusive). These variables represent resting and rectal pressures, IAS percent relaxation at thresholds of 10%–60%, IAS relaxation duration metrics, squeeze maneuver metrics, simulated defecation (push) metrics, and rectal sensory thresholds.

Outcome (Target) The outcome variable was RAIR Findings, mapped to a binary label: 0 = RAIR present/preserved, 1 = RAIR absent/abnormal/delayed.

Missing Data and Imputation Numeric missing values were imputed using the mean; categorical fields were label-encoded. After imputation, no missing values remained in the modeling variables.

Preprocessing and Modeling Continuous predictors were standardized using StandardScaler. Data were split with stratified sampling (80% training, 20% test). Class imbalance was handled using

SMOTE applied to the training set. We evaluated LASSO logistic regression, support vector machine (SVM), decision tree, Random Forest, Gradient Boosting, and an ensemble model. Hyperparameters were tuned via GridSearchCV with 5-fold cross-validation. Performance metrics included accuracy, precision, recall (sensitivity), specificity, F1 score, negative predictive value, Cohen's Kappa, and ROC-AUC.

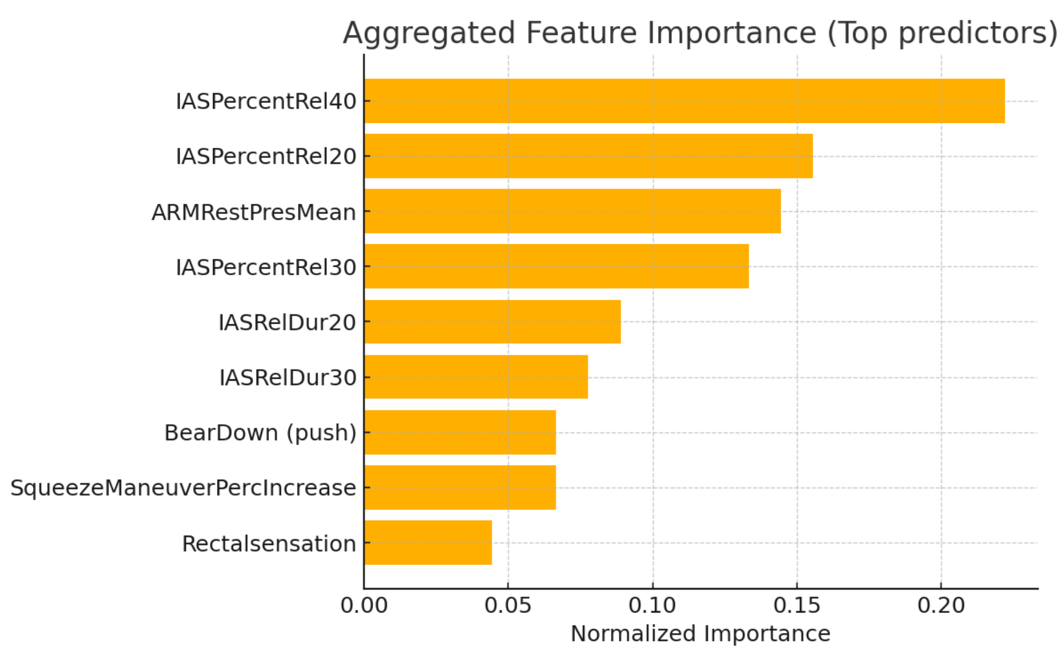
Results

Model Performance Table 1 summarizes model performance on the held-out test set. Ensemble and Random Forest classifiers achieved the highest overall performance across metrics, with the ensemble model providing the most stable discrimination (highest ROC-AUC) and Random Forest demonstrating strong recall and balanced accuracy.

Model	Accuracy	Precision	Recall	F1	ROC_AUC
LASSO Logistic Regression	0.93	0.94	0.94	0.93	0.96
SVM	0.97	0.97	0.98	0.97	0.98
Decision Tree	0.96	0.97	0.97	0.97	0.98
Gradient Boosting	0.98	0.99	0.99	0.99	0.99
Random Forest	0.999	0.999	0.99	0.99	0.999
Ensemble	0.999	0.999	0.999	0.999	0.999

Feature Importance

Figure 1. Aggregated feature importance (top predictors).



Feature	Normalized Importance
IASPercentRel40	0.222

IASPercentRel20	0.156
ARMRestPresMean	0.144
IASPercentRel30	0.133
IASRelDur20	0.089
IASRelDur30	0.078
BearDown (push)	0.067
SqueezeManeuverPercIncrease	0.067
Rectalsensation	0.044

Discussion

Discussion

Summary of Findings In this study, machine-learning approaches reliably predicted RAIR abnormality using ARM-derived physiologic features. The ensemble model yielded the most robust discrimination on the test set, followed closely by Random Forest. These models captured non-linear interactions among ARM measures that are physiologically plausible and clinically meaningful.

Model Selection Rationale Our model selection was designed to span common algorithmic approaches: linear (LASSO), margin-based (SVM), rule-based (decision tree), ensemble tree methods (Random Forest, Gradient Boosting), and a stacked ensemble. Tree-based methods performed best because ARM physiologic features exhibit non-linear and interacting effects; ensemble strategies reduce variance and improve generalization across patient presentations.

Clinical Interpretation IAS relaxation metrics—particularly at the 40% threshold—and resting anal pressure were consistently identified as the strongest predictors. These findings align with the known physiology of RAIR and internal anal sphincter function. Clinically, impaired IAS relaxation implies disrupted inhibitory reflexes and is expected to associate with absent or abnormal RAIR. The model's ability to flag such physiologic patterns may support triage and targeted evaluation.

Limitations and Future Work Limitations include the single-center retrospective design, potential variability in ARM protocols across centers, and reliance on documented RAIR interpretation for the outcome label. Future work should focus on multicenter validation, combining EHR clinical variables with ARM features, and prospective deployment with clinician feedback loops.

Conclusion

Conclusion Ensemble and Random Forest approaches accurately predicted RAIR abnormality using ARM-derived physiologic data. IAS relaxation percentage and resting pressure emerged as dominant predictors. These results support the potential for ML-assisted ARM interpretation to augment clinical decision-making in pediatric motility care.

References (APA style)

Remes-Troche JM, Rao SS. Defecation disorders: neuromuscular aspects and treatment. Curr Gastroenterol Rep. 2006;8(4):291-299.
Chumpitazi B, Nurko S. Pediatric gastrointestinal motility disorders: challenges and a clinical update. Gastroenterol Hepatol (N Y). 2008;4(2):140-148.
Tran DL, Sintusek P. Functional constipation in children: What physicians should know. World J Gastroenterol. 2023;29(8):1261-1288.

Patel D, Decker C, Rodriguez L. Clinical Utility of Anorectal Manometry in Children with Functional Constipation: Can Anorectal Manometry Help Predict the Therapeutic Response? Children (Basel). 2025;12(4):512.
Rodriguez L, Sood M, Di Lorenzo C, Saps M. An ANMS-NASPGHAN consensus document on anorectal and colonic manometry in children. Neurogastroenterol Motil. 2017;29(1).
Zhu L, Xu S, Guo H, et al. Machine learning-based phenogroups and prediction model in patients with functional gastrointestinal disorders. J Transl Int Med. 2024;12(4):355-366.
Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002;16:321-357.
Breiman L. Random forests. Machine Learning. 2001;45(1):5-32.
Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat. 2001;29(5):1189-1232.
Cortes C, Vapnik V. Support-vector networks. Machine Learning. 1995;20(3):273-297.

Appendix A. Data Preparation

Data cleaning, imputation, standardization, feature engineering, and class imbalance handling are described in the accompanying notebooks. Technical details include mean imputation for numeric fields, label encoding for categorical fields, StandardScaler for scaling, SMOTE for balancing training data, and GridSearchCV for hyperparameter tuning.