ARTIFICIAL INTELLIGENCE-BASED ANALYSIS OF UROFLOWMETRY PATTERNS IN CHILDREN: A MACHINE LEARNING PERSPECTIVE

Arslan F[1], Algorabi O[2], Ozkan OC[1], Turkan YS[2], Namli E[2], Genc YE[1], Yucel S[1] ,Cam K[1], Tarcan T[1,3]

1. Department of Urology, School of Medicine, Marmara University, Istanbul, Turkey
2. Department of Industrial Engineering, Istanbul University, Istanbul, Turkey
3. Department of Urology, School of Medicine, Koç University, Istanbul, Turkey

www.ics-eus.org/2025/abstract/550

Background

Uroflowmetry is a widely used non-invasive test for evaluating children with LUTS, though expert interpretation shows low agreement. This study investigates the potential of machine learning models to enhance the interpretation of uroflowmetry patterns.

Table 1. Performance of machine learning models

Machine Learning Models	Accuracy Rate%	Macro Average			Weighted Average		
		Precision	Recall	F1-score	Precision	Recall	F1-score
Decision Tree	81.80±1.47	0.81	0.80	0.80	0.82	0.82	0.82
Random Forest	84.80±1.33	0.85	0.82	0.83	0.85	0.85	0.85
CatBoost	84.80±1.47	0.84	0.82	0.83	0.85	0.85	0.85
XGBoost	85.00±2.90	0.85	0.82	0.83	0.85	0.85	0.85
LightGBM	83.00±2.10	0.83	0.79	0.80	0.83	0.83	0.83

Methods

Study Groups: Uroflowmetry tests from children aged 4–17 with LUTS were analyzed.
Three pediatric urology experts independently interpreted the patterns, resolving discrepancies by consensus.

Processing: Voiding parameters and flow rates at 0.5-second intervals were numerically processed.
80% of the data was used for training and 20% for testing across 5 machine learning classification models.

Results

Study population: 500 tests; 221 boys, 279 girls
Mean age: 9.17 ± 3.41 years

Observer agreement: Identical interpretations: 311 tests (62.2%) Different interpretations: 189 tests (37.8%), Fleiss' Kappa = 0.608

Training set patterns: Bell-shaped: 50.6%, Staccato: 20.6%, Tower: 10.4%, Plateau: 10.4%, Intermittent: 8%

Model performance: Highest accuracy: XGBoost (85.00% ± 2.90), Lowest accuracy: Decision Tree (81.80% ± 1.47)

Pattern classification accuracy: Highest: Intermittent (95–100%), Lowest: Tower & Plateau (61.54–73.08%)

Table 2. The models with the highest and lowest accuracy rates for all voiding patterns and their accuracy percentages

Voiding Patterns	Highest Accuracy	Accuracy Rate (%)	Lowest Accuracy	Accuracy Rate (%)
Bell Shaped	XGBoost	90.91	Decision Tree	85.77
Tower	Random Forest	73.08	Decision Tree XGBoost LightGBM	63.46
Staccato	CatBoost	83.50	LightGBM	79.61
Interrupted	XGBoost LightGBM	100	Random Forest	95
Plateau	XGBoost	71.15	LightGBM	61.54

Implications: The current trial demonstrated, for the first time, that machine learning models achieved a high accuracy rate in interpreting uroflowmetry patterns in children. Consequently, AI models have the potential to standardize the analysis of uroflowmetry voiding patterns in the future.