

# Utilizing machine learning to predict hospital admissions for paediatric COVID-19 patients

Liew Chuin-Hen<sup>1,2</sup>, Ong Song-Quan<sup>3</sup>, Ng Chun-Ern David<sup>4</sup>

<sup>1</sup>Hospital Tuanku Ampuan Najihah, Negeri Sembilan, <sup>2</sup>Clinical Research Centre of Hospital Tuanku Ampuan Najihah, Negeri Sembilan, <sup>3</sup>Institute For Tropical Biology and Conservation, University Malaysia Sabah, Sabah, <sup>4</sup>Hospital Tuanku Ja'afar, Negeri Sembilan

## ABSTRACT

**Introduction:** The COVID-19 pandemic has overwhelmed the healthcare systems globally. To cope with the high admission rate, it is crucial to identify COVID-19 patients who truly require hospitalization. Nevertheless, an accurate machine learning (ML) model to predict hospitalization in Asian children is lacking. This study aimed to bridge this gap by developing and validating ML models to predict hospitalization for children with COVID-19. **Methods:** We employed a cross-sectional design and included Malaysian children aged 0 to 12 with COVID-19 diagnosed between 1st February 2020 and 31st March 2022. The cohort was partitioned into training and validation groups. Feature selection was performed using the Recursive Feature Elimination (RFE) algorithm, and 7 classifiers were trained. Hyperparameter optimization was achieved using Grid Search. **Results:** We analyzed 1988 children with 29 study variables. The RFE identified 12 highly predictive variables for COVID-19 hospitalization (age, male sex, fever, cough, rhinorrhea, shortness of breath, vomiting, diarrhoea, seizures, body temperature, respiratory distress, and abnormal breath sounds). With external validation, Adaptive Boosting (AdaBoost) has the highest performance in predicting hospitalization (AUC=0.95, sensitivity=0.81, specificity=0.9, positive predictive value=0.89, negative predictive value=0.81). **Conclusion:** We showed that AdaBoost could classify the medical needs of pediatric COVID-19 (outpatient vs. hospital care) with high discriminative ability. It has the potential to serve as a Clinical Decision Support System, empowering frontline clinicians to make timely decisions using readily available clinical information without the need for expensive and time-consuming investigations. Ultimately, it may aid the hospital in conserving valuable hospital resources for future COVID-19 outbreaks.