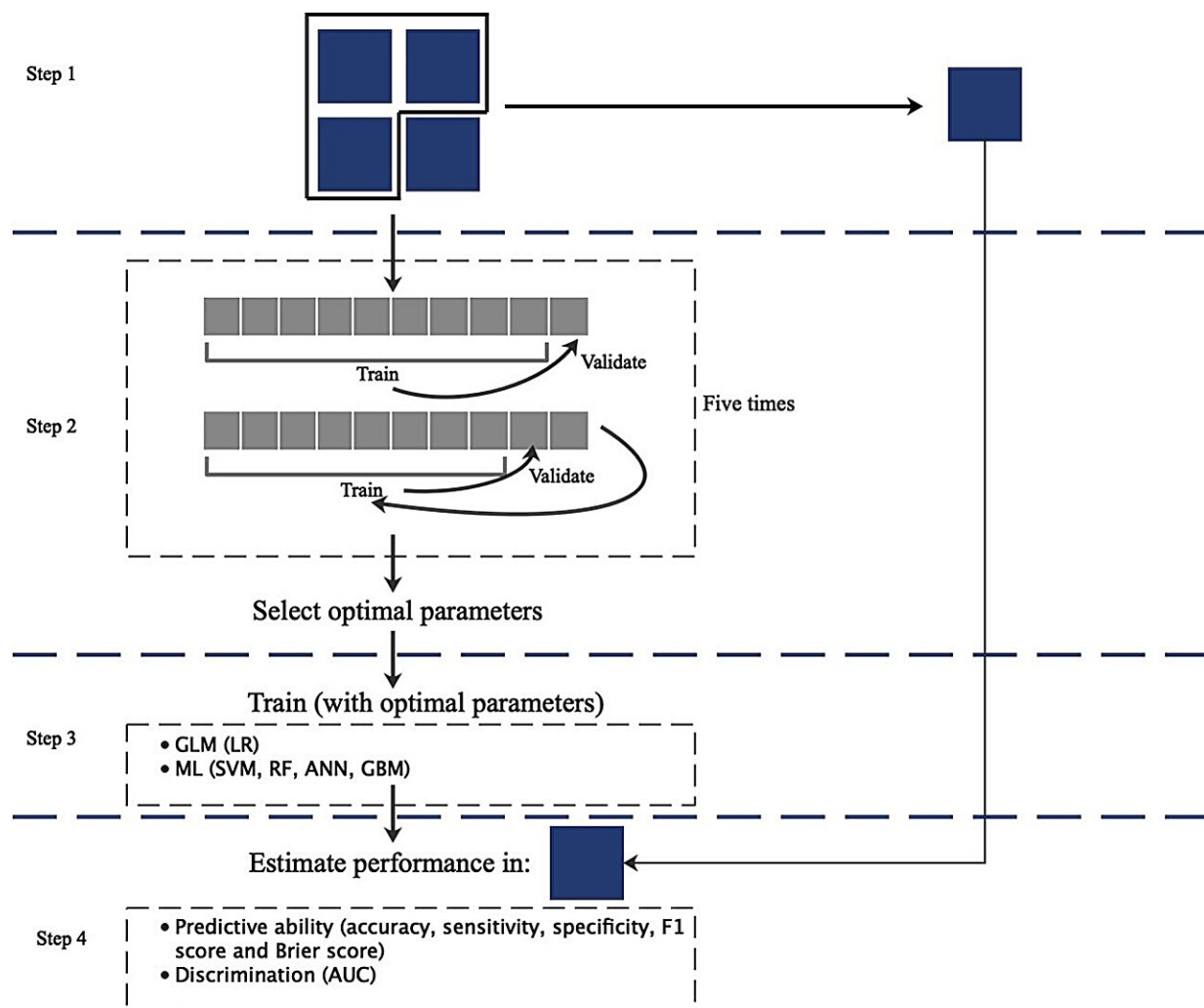


# Regression approach outperforms ML algorithms in predicting optimal surgical method in submucosal tumor patients

February 28 2024



Step 1 was selecting the training cohort (n = 184) and the internal validation cohort (n = 62) by random seeds; Step 2 was selecting the optimal parameters

through five times ten-fold cross-validation; Step 3 was the training of the final model with optimal parameters on the entire training data by the LR and ML ; Step 4 was validating the models of step 3 with the predictive ability and discrimination. GLM, Generalized linear model; LR, Logistic regression; ML, Machine learning; SVM, Support vector machines; RF, Random forests; ANN, Artificial neural networks; GBM, Gradient boosting machines; AUC, Areas under the receiver-operator characteristic curve. Credit: Zi-Han Geng, et al.

Submucosal tumors (SMTs) are usually found in the stomach and esophagus during an upper endoscopy. Submucosal tunneling endoscopic resection (STER) and non-tunneling endoscopic resection (NTER) are the two most commonly used techniques in the treatment of gastric and esophageal SMTs.

As novel technologies continue to shape the medical landscape, [machine learning](#) (ML) algorithms find increased application, demonstrating enhanced performance in various fields. Although some studies have evaluated the incremental value of flexible ML methods, comparisons with traditional logistic regression (LR) models are lacking.

To this end, a recent study by a team of researchers from China [published](#) in the journal *Gastroenterology & Endoscopy*, compared traditional regression models and ML algorithms to predict which technique performs better in surgery for submucosal tumors of the cardia.

Using key baseline predictive factors, ML algorithms and LR were conducted in 246 patients. For the ML algorithms, gradient-boosting machines, artificial neural networks, random forests, and support vector machines, were included.

For small sample-sized data, a technique for k-fold cross-validation was

exploited to avoid over-fitting. Meanwhile, the researchers tuned the parameters through several replications. Consequently, they quantified the discrimination (area under the curve, AUC) and predictive ability (Brier score, F1 score, specificity, sensitivity, and accuracy) of models.

"Four experts who have broad experience in STER and NTER in the upper GI tract (>1,000 cases) were asked to decide on the surgical technique for each patient. Predictors include mucosal status, growth pattern, maximum diameter, layer of origin, location, and morphology. Missing data were filled by Multiple Imputations by Chain Equations (MICE)," explained Quan-Lin Li, corresponding of the study.

The team found that LR outperformed among all groups (Brier score = 0.1398, F1 score = 0.7391, AUC = 0.8729, and predictive accuracy = 80.65 %). Morphology ranked in the top tier of all importance score lists, being the highest contributor to prediction accuracy. The direction of the gastroscope was also a key factor in most models. The other seven variables showed varying importance across different models.

"A limitation of our study is that the predictor used is relatively small, which potentially limited the performance of ML algorithms. Predictors with a higher correlation should be explored to improve ML algorithms. Besides, external validation is essential before applying prediction algorithms in [clinical practice](#), and our study did not include external validation cohorts because of the difficulty in generalizing inconsistent clinical settings from other centers," noted Li.

"The traditional regression approach outperformed ML algorithms for the prediction of the best surgical method in patients with SMTs. Further research is needed to validate and generalize our findings," concluded Li.

**More information:** Zi-Han Geng et al, A comparative analysis of prognostic regression models and machine learning algorithms in

surgical decision-making of cardial submucosal tumors, *Gastroenterology & Endoscopy* (2024). [DOI: 10.1016/j.gande.2023.12.001](https://doi.org/10.1016/j.gande.2023.12.001)

Provided by KeAi Communications Co., Ltd.

Citation: Regression approach outperforms ML algorithms in predicting optimal surgical method in submucosal tumor patients (2024, February 28) retrieved 11 May 2024 from <https://medicalxpress.com/news/2024-02-regression-approach-outperforms-ml-algorithms.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.