MLKD 2024
The First International Conference on
Machine Learning and Knowledge Discovery
Amirkabir University of Technology, December 18-19, 2024

Is Evaluation Based on Accuracy of Classification Algorithms Misleading? An Approach to Model Validation Using Bayes Error Rate

Hossein Kazemzadeh G [*] Adel Mohammadpour[†]

Abstract

Researchers have long regarded model accuracy as a tool for evaluating the performance of classification algorithms. The current evaluation approach, which relies solely on model accuracy, often leads to inappropriate evaluation of classifiers, regardless of the dataset's separability and complexity. This limitation underscores the need for a new, more comprehensive method. We claim that evaluation based on accuracy can be misleading, even if data separability or other complexity measure(s) are considered. We compare the performance of well-known classifiers on datasets generated from Gaussian models. We show that most algorithms' accuracy is greater than that of the best theoretical method, leading to overestimation. We label a model as invalid when its error dominates Bayes error. We introduce a procedure for finding invalid models and propose an algorithm for model validation based on the Bayes error rate.

Keywords: Classification, Evaluation, Validation, Golden Accuracy, Discriminant Analysis and Complexity Measure

[*]Department of Statistics, Faculty of Mathematics and Computer Science, Amirkabir University of Technology, kazemzadeh_hk@aut.ac.ir

[†]Department of Statistics, Faculty of Mathematics and Computer Science, Amirkabir University of Technology, Corresponding author:adel@aut.ac.ir