



# Local resampling for locally weighted Naïve Bayes in imbalanced data

Fatih Sağlam<sup>1</sup> · Mehmet Ali Cengiz<sup>1</sup>

Received: 22 December 2021 / Accepted: 19 August 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2023

## Abstract

Locally Weighted Naïve Bayes (LWNB) method establishes a weighted Naïve Bayes model in different neighborhoods of each query point. LWNB, like other classification methods, is affected by class imbalance. The class imbalance problem is the case where the class variable has a skewed distribution and causes the classification algorithms to be biased towards the majority class. It is possible to overcome this problem with resampling approaches such as undersampling and oversampling. Resampling on the data set may not reflect correctly on local regions, since regions are assumed to be independent of outside. Therefore, local regions should be considered without outside interference. In this study, we proposed a novel resampling approach that is applicable for both undersampling and oversampling. We examined how the imbalance of the data set should be reflected in each local region and aimed to prevent the imbalance problem by resampling data in the local regions separately. In this method, we calculated the appropriate resampling rate and the number of neighbors for each local region based on the data imbalance rate and the resampling rate which can be decided by the researcher. The proposed approach was compared with the classical resampling approaches on 25 datasets that are frequently used in the literature and achieved promising results.

**Keywords** Locally weighted learning · Class imbalance · Resampling · Oversampling · Undersampling · SMOTE

**Mathematics Subject Classification** 62H30 · 62D99

---

✉ Fatih Sağlam  
fatih.saglam@omu.edu.tr

Mehmet Ali Cengiz  
macengiz@omu.edu.tr

<sup>1</sup> Department of Statistics, Faculty of Sciences, Ondokuz Mayıs University, 55139 Samsun, Turkey