

Modified Leader Algorithm for Under-Sampling the Imbalanced Dataset for Classification

[S. Karthikeyan](#) & [T. Kathirvalavakumar](#)

Conference paper | [First Online: 27 August 2021](#)

692 Accesses

Part of the [Lecture Notes in Networks and Systems](#) book series (LNNS, volume 213)

Abstract

Data classification with a standard classifier automates the manual classification process in many fields. In a two-class dataset, when the number of samples in one class is more in number than the other class, namely, imbalanced, then the performance of a classifier gets degraded due to the limited availability of the training instances in a particular class. To overcome the problems with the imbalanced datasets, a new under-sampling method has been proposed with the baseline idea of an incremental clustering technique. Clusters are formed from the sum of features of the instances instead of finding distance between patterns. Representatives of the clusters are average of the instances of the cluster. Proposed algorithm has the ability to solve the problems than the existing under-sampling approaches with k-means algorithm and leader algorithm. The results produced through the proposed algorithm work better during the classification with good accuracy and reduced misclassification rate in both major and minor classes.

Keywords

[Imbalanced data](#)

[Under-sampling](#)

[Incremental clustering](#)

[Leader algorithm](#)

Access via your institution



Chapter

EUR 29.95

Price includes VAT (India)

- Available as PDF
- Read on any device
- Instant download
- Own it forever

Buy Chapter

eBook

EUR 192.59

Softcover Book

EUR 229.99

Tax calculation will be finalised at checkout

Purchases are for personal use only

[Learn about institutional subscriptions](#)

Sections

References

[Abstract](#)

[References](#)

[Author information](#)

[Editor information](#)

[Rights and permissions](#)

[Copyright information](#)

[About this paper](#)