ABSTRACT

Drug Target Interaction (DTI) plays a vital role in pharmaceutical research by facilitating drug discovery, therapeutic repositioning, and identification of side effects. Computational techniques, particularly machine learning, are increasingly used due to cost and time efficiencies. However, DTI prediction faces challenges, notably class imbalance. This study investigates the application of oversampling followed by Ant Colony Optimization (ACO) based undersampling, to address class imbalance problem in DTI prediction using Sequential Hybrid Sampling (SHS) approach. Several oversampling methods such as SMOTE, Random Oversampling (ROS), and ADASYN were used in the experiment to get the best results. The evaluation results show that by using F1-score as the fitness function, it is found that using ACO as undersampling based with SHS approach can improve the performance of the classifier. Gradient boosting is used as the classification method and F1-score, G-Mean, and Balanced Accuracy are used as evaluation metrics. The result show that implementation SHS with SMOTE and ACO were achieved the highest F1-score of 43.24%, Geometric Mean (G-Mean) of 50.75%, and Balanced Accuracy Score (BAS) of 62.41%. These results highlight that the implementation of SHS with ACO can improve the performance of the classifier in DTI prediction compared to the classic oversampling technique. By advancing the understanding and methodologies for handling class imbalance in DTI prediction, this study contributes to the broader goal of enhancing drug discovery and development processes.

Keywords: DTI, SHS, Oversampling, Undersampling, ACO, SMOTE