ABSTRACT

The drug and target interaction are essential in drug discovery to know how the drug interacts with the target. The identification process needs laboratory experiment, which makes a wasted cost. Recently, an alternative approach is a computational method using machine learning. This method can help to identify the prediction of drug and target with the assumption each drug has a similar target. The drugs cannot interact with all the targets making in different number of interacting pairs and non-interacting pair make the data imbalanced. This imbalance issue can reduce the performance of the prediction. The balancing method, i.e., under-sampling, oversampling, and hybrid-sampling, is used to handle the imbalance issue. Under-sampling can help to remove the majority class until its balanced, but this method drawback can remove the potential majority class. To overcome the drawback of under-sampling, a combination of under-sampling and optimization algorithms, such as Firefly Algorithm. The Firefly Algorithm Under-sampling to handle the imbalance issue in drug target interaction prediction. This research aimed to implement Firefly Algorithm Under-sampling (FAUS) to handle the imbalance issue in drug-target interaction prediction. We conducted the Firefly algorithm under-sampling with the variation of Alpha, Beta, and Gamma Parameters. These parameter variations significantly impact the optimization process and contribute to achieving the best performance. Based on the results, we found that FAUS gives the best value, with accuracy and F1 scores of 0.94 and 0.52, respectively.

Keywords: *DTI, Firefly Algorithm, Under-sampling*