

Abstract– Social media has emerged as immensely popular and favored platforms among the masses today. Twitter, being one of the most renowned social media platforms, allows users to express themselves through tweet postings. Retweeting is a crucial feature on Twitter, enabling users to disseminate tweets authored by others. In this context, this research aims to predict retweet behavior using User-Based, Content-Based, and Time-Based features, coupled with an Artificial Neural Network classifier optimized with Grey Wolf Optimization. One of the challenges in retweet prediction lies in class imbalance, where the number of retweets on certain tweets is significantly disproportionate compared to others. To address this issue, this study implements undersampling and oversampling techniques. Undersampling reduces the number of samples from the majority class, whereas oversampling involves duplicating or synthesizing samples from the minority class, thereby creating class balance. The research successfully achieves promising results in retweet prediction. After applying oversampling techniques, the classification process attains an accuracy of 85.58%, precision of 87.77%, recall of 83.92%, and F1-score of 85.80%. These results demonstrate the effectiveness of the proposed method in retweet prediction and handling class imbalance issues.

Keywords: Social Media; Twitter; retweet; Prediction; Imbalance Class; undersampling; oversampling; Artificial Neural Network; GreyWolf.