IMPROVING ACCURACY METRIC WITH PRECISION AND RECALL METRICS FOR OPTIMIZING STOCHASTIC CLASSIFIER

Hossin M.[1,2], Sulaiman M.N.[1], Mustpaha N.[1], Rahmat R.W.[1]

[1]*Universiti Putra Malaysia (UPM), Malaysia, mhossin78@gmail.com,*
{nasir, norwati, rahmita}@fsktm.upm.edu.my
[2]*Universiti Malaysia Sarawak (UNIMAS), Malaysia, hmohamma@fcs.unimas.my*

ABSTRACT. All stochastic classifiers attempt to improve their classification performance by constructing an optimized classifier. Typically, all of stochastic classification algorithms employ accuracy metric to discriminate an optimal solution. However, the use of accuracy metric could lead the solution towards the sub-optimal solution due less discriminating power. Moreover, the accuracy metric also unable to perform optimally when dealing with imbalanced class distribution. In this study, we propose a new evaluation metric that combines accuracy metric with the extended precision and recall metrics to negate these detrimental effects. We refer the new evaluation metric as optimized accuracy with recall-precision (OARP). This paper demonstrates that the OARP metric is more discriminating than the accuracy metric and able to perform optimally when dealing with imbalanced class distribution using one simple counter-example. We also demonstrate empirically that a naïve stochastic classification algorithm, which is Monte Carlo Sampling (MCS) algorithm trained with the OARP metric, is able to obtain better predictive results than the one trained with the accuracy and F-Measure metrics. Additionally, the *t*-test analysis also shows a clear advantage of the MCS model trained with the OARP metric over the two selected metrics for almost five medical data sets.

INTRODUCTION

Instance selection (IS) is one of the classification methods which aim to reduce the instances as much as possible and simultaneously attempt to achieve the highest possible classification accuracy. From the previous studies, some of the IS methods are developed using stochastic methods such as Monte Carlo (Skalak, 1994), genetic algorithm (Garcia-Pedrajas et al., 2010) and tabu search (Ceveron & Ferri, 2001). In general, these algorithms use the training stage learns from the data and at the same time attempt to optimize the solution by discriminating the optimal solution from the large space of solutions. In order to find the optimal solution, the selection of suitable evaluation metric is essential. According to Ranawana and Palade (2006), to select the suitable evaluation metric for discriminating an optimal solution, the selected evaluation metric must be able to maximize the total number of correct predicted instances in every class. In certain situation, it is hard to build an optimized classifier that can obtain the maximal value for every class. This is because, traditionally, most of the stochastic classification algorithms employ the accuracy rate or the error rate (1-*accuracy*) to discriminate and to select the optimal solution. In (Huang & Ling, 2005; Ranawana & Palade, 2006; Wilson, 1996), they have demonstrated that the simplicity of this accuracy metric could lead to the sub-optimal solutions. For instance, when dealing with imbalanced class instances, it is often happen that the classification model is able to perform extremely well on a large class