Source: <https://dl.acm.org/doi/pdf/10.5555/2107653.2107659>

Approaches measurements

1. the geometric mean

Separate accuracy measures on the two classes (a1 and a2)

√ a1 × a2: geo mean of the correctly classified proportions of classes 1 and 2

1. average precision and recall, or F-measure

Solutions:

1. balancing the training set using under or over sampling
2. the SMOTE (Chawla et al., 2004) algorithm is a technique for creating new instances of the minority class,

FEATURE SELECTION STUFF:

1. IG: information gain
2. Mutual information
3. Symmetric uncertainty : https://www.aclweb.org/anthology/R13-1089.pdf