

IGARSS 2015, Milan, Italy, 26-31 July 2015

Proper comparison among methods using a confusion matrix

^{1,2} B.P. Salmon, ^{2,3} W. Kleynhans, ^{2,3} C.P. Schwegmann and ¹ J.C. Olivier

¹ School of Engineering and ICT, University of Tasmania, Australia

² Remote Sensing Research Unit, Meraka Institute, CSIR, South Africa

³ Department of Electrical, Electronic and Computer Engineering, University of Pretoria, South Africa. bpsalmon@utas.edu.au

Abstract

An important aspect of research in the remote sensing field is to objectively compare different classifiers. This is the foundation of hundreds of research projects and in this paper we will address some raising concerns when evaluating solutions for classification of data sets with skewed class distributions. The quality of assessment is based on the problem specified by the user and the corresponding hypothesis defined. This hypothesis will determine how two or more classifiers are scored to determine which one is better for a particular application. In this paper we present two experiments that illustrate how, if unaware and misunderstood, statistical measurements can be misleading. One experiment is based on a Synthetic Aperture Radar image with a highly skewed class distribution and the second experiment is based on a Landsat image with a minor skewed distribution. From both experiments it can be seen that ill-defining the problem, can lead to false statements and the reporting of statistically invalid conclusions.