

Title:**Detecting outliers in credit scoring using logistic regression****Authors:**

T de la Rey & JH Venter
Centre for Business Mathematics and Informatics
North-West University
South Africa

Abstract:

In credit scoring logistic regression remains an old favourite. The estimated logistic regression (LR) relationship may be severely affected by outliers; moreover outlying individuals are atypical in credit scoring which makes it important to identify them. This motivates the need for robust logistic regression combined with outlier detection. In this talk we introduce a procedure that achieves these goals using maximum likelihood in conjunction with downweighting. The procedure selects two sets of weights, namely high and low weights and then splits the data optimally into two subsets to which the high and the low weights are attached, the subset with the low weights containing the observations that are more likely to be outliers. A corresponding weighted ML estimator of the regression coefficients is computed. This is used to estimate the response probabilities of the individual observations. Observations with a “good” response but low probability for this response and observations with a “bad” response but high probability for this response can then be classified as outliers. The procedure is illustrated with relevant examples.