

1086-62-2417

Xin Zhang, Jun Li, Vance Wong, Brian Noland and Daniel Jeske*

(daniel.jeske@ucr.edu), Department of Statistics, 1340 Olmsted Hall, Riverside, CA 92521.

Logistic Regression Classifiers with Longitudinal Data.

We propose a robust two-group classification methodology for longitudinal data applications. The classifier is built as follows. A nonparametric mixed effects (NME) model is fit to training data. Nonparametric basis functions are used to capture the underlying group-dependent time trends and subject-specific random intercepts and slopes are used to tune the model for subject heterogeneity. The NME is used to obtain empirical best linear unbiased predictions (eBLUPs) for the slopes and intercept of each subject. A logistic regression model whose explanatory variables are eBLUPs is fit and used to evaluate the posterior membership probability for, say Group 1. A classification threshold for the posterior membership probability is determined from a receiving operating curve (ROC) analysis. New subjects are classified by evaluating their eBLUPs, inserting them into the logistic regression model, and comparing the posterior membership probability for Group 1 to the selected threshold. We illustrate the use of this methodology and then compare and contrast it's effectiveness to an alternative approach where the measured responses at each time point are used directly in a logistic regression classifier. (Received September 25, 2012)