A NOVEL DATA MINING TECHNIQUE WITH REVERSE ENGINEERING CAPABILITY: APPLICATION TO AUTOMATED DETECTION OF FLUX TRANSFER EVENTS

H. Karimabadi[1,2], T. Sipes[2], Y. Wang[3], and B. Lavraud[4]

[1]*SciberQuest, Inc.,* [2]*UCSD,* [3]*Goddard Space Flight Center,* [4]*Los Alamos National Laboratory*

A novel data mining method called MineTool is introduced which, by virtue of automating the modeling process and model evaluations, makes it more accessible to nonexperts. In all cases tested, the algorithm outperformed in terms of both accuracy and speed the standard techniques such as artificial neutral net, genetic algorithm, regression trees and support vector machines. The technique aggregates the various stages of model building into a four-step process consisting of (1) data segmentation and sampling, (2) variable preselection and transform generation, (3) predictive model estimation and validation, and (4) final model testing. Optimal strategies are chosen for each modeling step. However, the modular design of the MineTool enables the substitution of alternative strategies in any of the four modeling steps. A notable feature of the technique is that the final model is always in closed analytical form rather than black box form of most other techniques. MineTool can be used for analysis of data (e.g., time series) as well as images. The utility of the technique is illustrated through an application to automated detection of Flux Transfer Events in the Cluster data. Flux Transfer Events are thought to be flux ropes associated with the collisionless magnetic reconnection process at the Earths dayside magnetopause.