Earth **Surface**
Dynamics
Discussions

Interactive comment on "An introduction to learning algorithms and potential applications in geomorphometry and earth surface dynamics" *by* Andrew Valentine and Lara Kalnins

J. Becker

loveforlife.jj@gmail.com

Received and published: 17 March 2016

Page 9 line 5

As the 3rd author of Martin 2015, and the main author and "debugger" of the code used there, I wanted to throw in my two cents worth about how important it is to normalize both the "features" (aka predictors) and the "response" (aka measurements).

This doesn't need to be complex, simply "Student-ize" by removing the mean of every column of the training data, (store that number), and dividing the detrended data by it's standard deviation, (also store that).

When making a prediction after training, it is important to apply the mean and std dev

of the TRAINING data to the data used in the PREDICTION.

Often "software packages" do this automatically (aka Random Forests in MATLAB), but other times this is not the case and it will dramatically hurt the prediction. For example, the recent advances in "deep learning" for Machine Vision were only achieved after the researchers understood they needed to subtract the "average image" obtained from all the training images -BEFORE- training and prediction...

-jj

Interactive comment on Earth Surf. Dynam. Discuss., doi:10.5194/esurf-2016-6, 2016.