
**INVESTIGATING THE EFFECT OF DATA NORMALIZATION ON
PREDICTIVE MODELS**

Ajiboye, Adeleke Raheem

Department of Computer Science,
University of Ilorin,
Ilorin, Nigeria
ajibraheem@live.com

Ajiboye, Ismail Kolawole

Department of Computer Science
Abdul-Raheem College of Advanced Studies
Igbaja, Nigeria.
fabik2000@gmail.com

Salihu, Shakirat Aderonke

Department of Computer Science
University of Ilorin
Ilorin, Nigeria.
shaksoft@yahoo.com

Rasheed Adekola Tomori

Computer Services & Information Technology
University of Ilorin
Ilorin, Nigeria
tomorirasheed@yahoo.com

Abstract:

The creation of predictive model using a supervised learning approach involves the task of building a model of the target variable as a function of the explanatory variables. Before a model is created, it is necessary to put the data in a suitable format. Studies have shown that normalization of data is crucial to descriptive mining as it improve the accuracy and efficiency of mining algorithms. However, in the case of prediction, it is not in all cases that predictive models are created from normalized data. This paper presents the experimental results of investigating the effect of normalizing the input variables on models created for prediction purposes. Experiments are conducted for the creation of predictive models from two different sets of equal size of data using neural network techniques. The trained network models created with the same architecture and configurations are subsequently simulated using a set of untrained data. The evaluation results and the comparison of the models created through the two data sets of different format reveals that, the model created from a normalized data appears to be more accurate as a decrease in error by 0.003 are consistently recorded. The model also converges much earlier than the model created from the data that does not undergo any form of normalization.

Keywords: data normalization, data pre-processing, predictive model, supervised learning.