

International Conference on **Big Data Analysis and Data Mining**

May 04-05, 2015 Kentucky, USA

Treatment of missing values in data mining

Waqas Ilyas and Syed Saeed Ur Rahman Shah
Superior University Lahore, Pakistan

Data Mining has pushed the realm of information technology beyond predictable limits. Data Mining has left its permanent marks on decision making in just in few years of its inception. Missing value is one of the major factor, which can render the obtain result beyond use attained from specific data set by applying Data Mining technique. There could be numerous reasons for missing values in a data set such as human error, hardware malfunction etc. It is imperative to tackle the labyrinth of Missing Values before applying any technique of Data Mining; otherwise, the information extracted from data set containing missing values will lead to the path of wrong decision making. There are several techniques available to control the issue of missing values such as replacing the missing value with: (a) Closest value, (b) mean value and (c) median value etc. Some algorithms are also used to deal with the problem of missing values such as k-nearest neighbour. In this paper, we will review certain techniques and algorithms to deal with the puzzle of Missing Values whereby achieving pure data set (i.e. data set without Missing Value) which in-turn will lead to path of correct and accurate decision making.

waqasilyas12@gmail.com

Notes: