SMART REACH

(S/W CREATORS & TRAINERS)

ISO 9001:2008 CERTIFIED COMPANY

Ph: 9585554590, 9585554599

Email: support@salemsmartreach.com

URL: www.salemsmartreach.com

Decision Trees for Mining Data Streams Based on the Gaussian Approximation

Abstract:

Since the Hoeffding tree algorithm was proposed in the literature, decision trees became one of the most popular tools for mining data streams. The key point of constructing the decision tree is to determine the best attribute to split the considered node. Several methods to solve this problem were presented so far. However, they are either wrongly mathematically justified or time-consuming. In this paper, we propose a new method which significantly outperforms the tree algorithm and has a solid mathematical basis. Our method ensures, with a high probability set by the user, that the best attribute chosen in the considered node using a finite data sample is the same as it would be in the case of the whole datastream.