Physical Data Warehouse Design Using Neural Network

Year of Publication: 2010

Authors:

Mayank Sharma
Navin Rajpal
B.V.R.Reddy

Abstract

Performance of the data warehouse depends on physical design. Index selection and storage of multidimensional data bases are important activities of physical designing process. Conventional indexing techniques such as bitmaps, B-trees and hash based indexing systems need large storage space for storing indexes along with data itself. Spelling variants, misspellings and transliteration differences are source of uncertainty in data with in the databases. Misspelled and distorted key values are also hard to map in present indexing systems. In this paper neural network based physical design is suggested, a class of artificial neural network known as self-organizing net is used for indexing data warehouse at physical level. Indexes of active neurons will be used for generating indexes for the data values. In conventional indexing techniques every key value is mapped to a specific point in space, while in neural network based database indexing system, every key value is mapped to a region in space. This region is a class to which the key values of similar type belong. Indexes generated through this method used optimal space for storage, as only final weight matrices after training of neurons are stored. Self-organizing net based indexing is very robust as distorted key values get indexed to right classes. Accuracy of our self-organizing net based indexing system in
mapping key values with distorted keys is found to be high.

Reference


http://dx.doi.org/10.1023/A:1009777418785


http://doi.acm.org/10.1145/119995.115807


Wisconsin-Madison, CS Department, May, 1996.

Index Terms

Computer Science

Databases

Key words

Self-organizing net

multidimensional databases

indexing