A SCALABLE MODEL FOR COLLABORATIVE FILTERING USING SELF-ORGANIZING MAPS AND EXTENDED DENSITY-BASED INSTANCE SELECTION

Jerome, Otobong B.
Obafemi Awolowo University
NIGERIA
jeromeblacq@gmail.com

ABSTRACT

Increase volume of input data poses a problem of scalability to traditional memory-based recommender systems, In this paper, we propose a recommendation system model that is both efficient and highly scalable. The model is also very intuitive which makes it easy to implement and debug on a large scale and commercially. This model is an ensemble of several algorithms. The recommendation here is done in two stages, in the first stage we build a condensed representation of the input space using self-organising maps to cluster the input then further extracting the most relevant instances based on their neighbourhood densities ordering. The final case base carries the overall topology of the original big-data but is considerably reduced in size. In the second stage, the online recommendation is done using the similarity of the user to the users in the condensed case base. We achieved a recommendation accuracy of 96.1%, at a very high speed online. Certain system constants play an important role in the system"s accuracy, parameters such as SOM matrix size, the neighbourhood function as well as the similarity metric, were determined here experimentally using different accuracy measures. This showed that the accuracy of recommendation depends more on the ability of the data set to capture the nature of the users" space than on the size of the data set. In most cases, the data set contains lots of redundant data points and as such it is useful to extract such points as have more information to represent the neighbourhood than to make use of the entire neighbourhood.

Keywords: Collaborative Filtering, Instance Selection, Recommendation.