

Une approche connexionniste pour l'extension de l'OLAP à des capacités de prédiction

Wiem Abdelbaki***, Sadok Ben Yahia*, Riadh Ben Messaoud**

*Faculté des Sciences de Tunis, Tunisie.

**Faculté des Sciences Économiques et de Gestion de Nabeul, Tunisie.

Résumé. Les outils de l'analyse en ligne (OLAP) permettent à l'utilisateur de réaliser des tâches exploratoires dans les cubes de données. Cependant, ils n'offrent aucun moyen pour la prédiction ou l'explication des faits. En vue de renforcer le processus de l'aide à la décision, plusieurs travaux ont proposé l'extension de l'analyse en ligne à des capacités plus avancées. Dans cet article, nous proposons une nouvelle approche d'extension de l'analyse en ligne à des capacités de prédiction à deux phases. La première est une phase de réduction des dimensions des cubes de données, qui repose sur l'analyse en composantes principales (ACP). La deuxième est une phase de prédiction dans laquelle nous introduisons une nouvelle architecture de perceptrons multicouches (PMC). Notre étude expérimentale a montré une capacité de prédiction prometteuse, ainsi qu'une bonne robustesse dans le cas d'un cube épars.

1 Introduction

Les entrepôts de données représentent une solution convenable au problème du stockage et de l'analyse des données historisées d'une entreprise (Inmon, 1996). En fait, ils permettent de stocker les données dans des structures multidimensionnelles appelées *cubes de données*. Ces cubes, contiennent des données, qui sont déjà consolidées et intégrées (Inmon, 1996). Par ailleurs, la visualisation d'un cube dévoile généralement une structure éparsée. En effet, les cellules vides reflètent des événements passés, qui n'ont pas eu lieu à l'instant présent ou des événements futurs, qui n'ont pas encore eu lieu.

Nous pensons qu'il serait très intéressant pour l'utilisateur de pouvoir anticiper les valeurs de la mesure d'une cellule vide pour une meilleure prise de décision. Par exemple, il serait très utile à une société de vente de voitures de prédire le chiffre d'affaire, que peut réaliser une nouvelle agence dans une nouvelle ville avant de lancer ce projet. Cet indicateur encouragera, ou non, l'investissement de cette société dans cette agence. En effet, l'information manquante, qui concerne une cellule, est potentiellement existante dans son voisinage. Cependant, la technologie OLAP n'offre aucun moyen pour l'explication ou la prédiction des faits.

En réponse à ce constat, plusieurs études proposent d'étendre l'OLAP pour répondre à différents objectifs comme l'exploration des cubes de données (Sarawagi et al., 1998) et l'extraction des règles d'association (Ben Messaoud, 2006). Récemment, de nouvelles recherches proposent d'étendre OLAP à des capacités de prédiction afin d'anticiper des événements futurs (Bodin-Niemczuk et al., 2008; Chen et al., 2005).