

Etudiant : NGUYEN Huu Khang - 21506865

Rendu Article 4 :

Extraction de Bases pour les Règles d'Association à partir des Itemsets Fermés Fréquents

1 - Résumé de l'article.

L'extraction des règles d'association est une étape importante dans l'analyse des relations cachées dans la base de données afin de mieux utiliser les données et appliquer dans les domaines différents. Cette étape consiste 4 différentes phases qui part de la préparation des données jusqu'à l'analyse des résultats. Dans cet article, l'auteur se concentre principalement sur deux problèmes concernant le temps d'exécution ainsi que la pertinence et l'utilité des règles d'association. Les solutions ont été proposées dans la suite de l'article. Concernant sur le problème du temps d'exécution, l'auteur présente les algorithmes efficaces qui a pour le but de réduire le nombre d'itemsets pour la suite de processus. Les solutions du deuxième problème consistent à réduire les règles d'association ou de sélectionner les sous-ensembles afin d'éviter tous les redondances possibles. L'auteur montre également l'utilisation des bases des règles d'associations non redondances minimales qui ne contiennent aucunes pertes d'informations ainsi que leur pertinence et leur support. L'article se termine sur l'interprétation des résultats obtenus sur les différents jeux des données.

2 - Préciser quelle est la différence entre des itemsets clos et des itemsets maximaux.

Tous les sur-ensembles des itemsets maximaux fréquents sont inféquents.

3 - A votre avis, est-ce que le nombre d'itemsets maximaux est supérieur ou inférieur au nombre d'itemsets clos ? Justifier la réponse.

Le nombre d'itemsets maximaux est inférieur à celles des itemsets fermés car d'après l'article, les itemsets maximaux sont aussi les itemsets

fermés alors que les itemsets fermés ne sont pas forcément les itemsets maximaux.

4 - Est il possible à partir des itemsets maximaux d'obtenir l'ensemble des itemsets clos ? Justifier la réponse.

Oui, c'est possible de retrouver l'ensemble des itemsets fermés car on peut retrouver tous les itemsets fréquents de l'ensemble des itemsets fermés ainsi que leur support. Cela nous permettra de retrouver les itemsets fermés en appliquant la fermeture de la connexion Galois qui consiste d'enlever tous les sous-ensembles non maximaux d'items.

5 - Dans le cadre de votre projet de classification vous avez utilisé différents traitements. Préciser comment l'utilisation de règles d'association pourrait aider à améliorer les classifieurs. Indications : les items peuvent être des mots pré-traités et la classe est connue.

Avec les règles d'association nous pouvons améliorer notre classifieur en appliquant les règles d'association sur les mots pré-traités ainsi que les classes connues dans la pré-traitement. Nous pourrions donc appliquer les différentes méthodes sur les différents critères afin d'améliorer la précision de notre classifieur.