

Détermination du niveau de consommation des abonnés en téléphonie mobile par la théorie des ensembles flous

Rachid EL MEZIANE (*), Ilham BERRADA (*), Ismail KASSOU (*), Karim BAINA (*)

Laboratoire Al Khawarizmi - ENSIAS - BP 713 - Agdal - Rabat - Maroc
(*){meziane, iberrada, kassou, baina@ensias.ma}

Résumé. La détermination du niveau de consommation chez les clients est essentielle pour tout objectif de segmentation stratégique et de churn. Nous présentons sur un cas réel l'utilisation de la théorie des ensembles flous pour la définition d'une fonction d'appartenance permettant d'évaluer, de manière précise, le niveau de consommation, des abonnés en téléphonie mobile.

1 Contexte

Notre travail s'inscrit dans le contexte d'un projet de fouille de données mis en oeuvre à Maroc Telecom et visant à mieux connaître la clientèle de la téléphonie mobile. Le niveau de consommation d'un abonné est souvent calculé à partir de la durée facturée qui s'avère insuffisante pour la plupart des cas. En effet, deux abonnés peuvent avoir la même durée d'appel pour des services différents mais sans avoir le même degré de consommation. D'où la nécessité d'introduire d'autres critères dans la détermination du niveau de consommation.

2 Problématique et approche de résolution préconisée

La problématique à laquelle on s'intéresse consiste à établir une échelle de mesure permettant de quantifier les niveaux de consommation afin discriminer entre les abonnés (Viertl, R. (2005)). L'approche de résolution proposée comporte trois étapes principales. Son originalité réside dans l'utilisation de la théorie des ensembles flous à travers la définition expérimentale d'une fonction d'appartenance (Mitaim, S. et B, Kosko. (2001)).

Dans une première étape, on attribue un score aux abonnés par rapport aux critères de type catégoriels (trafic, produits, services, plage horaire) caractérisant le niveau de consommation. La binarisation de chaque modalité de ces critères induit la création de plus de 60 variables indicatrices dans notre exemple qui traite 2 millions d'enregistrements. Afin de réduire la taille de ces indicatrices, l'Analyse des Correspondances Multiples (ACM) a été utilisée fournissant ainsi 10 facteurs expliquant 80,62% d'inertie totale.

L'objectif de l'étape 2 est la segmentation des abonnés par produits et services afin de discriminer entre les abonnés en se basant sur le comportement d'utilisation des produits et services. Les facteurs obtenus par l'ACM ont été utilisés comme variables d'entrée des différents algorithmes non supervisés (K-means, Two Step, Réseau de Kohonen) qui ont été comparés. Le réseau de Kohonen a été plus concluant en terme d'homogénéité entre les classes. Les inerties intra classes de chaque facteur ont ensuite été utilisées dans l'étape 3 comme indicateur de la variation du niveau de consommation au sein de chaque classe. Un tel indicateur a permis d'établir une mesure du niveau de consommation tenant compte de la durée facturée en appliquant la théorie des ensembles flous (Masson, M. H. (2003)).