

LE SCORING AVEC UNIWIN Plus

DESCRIPTIF

Le scoring est une technique permettant de classer des individus caractérisés par des variables descriptives qualitatives dans deux catégories préexistantes en leur affectant un score d'appartenance.

La première étape de cette technique consiste à transformer les variables descriptives qualitatives d'origine en coordonnées sur les axes factoriels issus d'une Analyse des Correspondances Multiples (ACM) des variables qualitatives.

La deuxième étape est une Analyse Factorielle Discriminante (AFD) de ces données, avec une possible sélection des axes factoriels significatifs de l'ACM.

La fonction discriminante obtenue est ensuite exprimée en fonction des indicatrices des modalités des variables qualitatives d'origine et les coefficients de la fonction de score sont évalués.

L'analyse se poursuit en calculant les scores des individus, la sensibilité, la spécificité, le niveau de qualité ROC et l'indice de Gini pour la courbe Lift.

Un rapport général de synthèse est proposé ainsi que des graphiques des coefficients de la fonction de score, des courbes de répartition (avec ou sans zone d'indécision), des courbes de densité, de la courbe ROC et de la courbe Lift.

Une option de simulation permet d'évaluer l'évolution du score d'un individu lorsque une ou plusieurs de ses modalités d'origine sont modifiées.

Des tableaux résumés et détaillés des classements de l'Analyse Factorielle Discriminante et du Scoring sont calculés et les principaux résultats du scoring peuvent être sauvegardés sur fichier.

L'étude d'une population nouvelle d'anonymes peut également être réalisée. Elle évalue les scores de ces nouveaux individus.

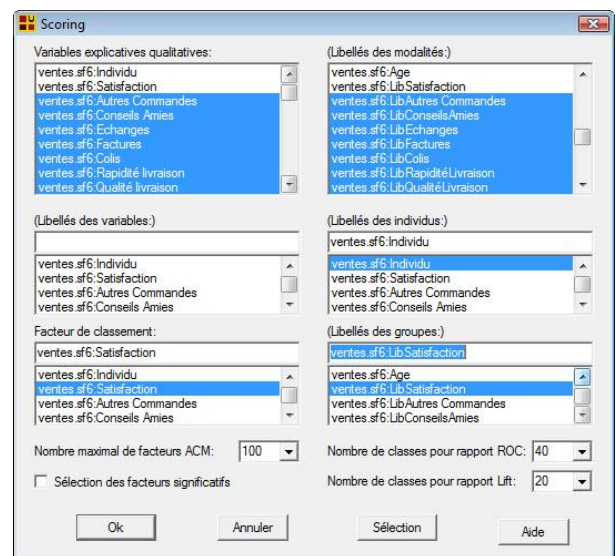
Les applications de cette technique sont multiples, notamment pour la caractérisation de demandeurs de crédit ou l'étude de la satisfaction des clients. Illustrons ce dernier cas.

EXEMPLE

Une enquête a été effectuée concernant 1.596 clientes d'une société de vente par correspondance. Les clientes ont été interrogées sur 22 thèmes :

1. Commander d'autres produits
2. Conseiller les produits à des amies
3. Gestion des retours/échanges
4. Exactitude des factures
5. Conformité du contenu du colis
6. Rapidité de la livraison

7. Qualité de la livraison
8. Diversité des modes de règlement
9. Flexibilité en cas d'annulation d'une commande
10. Facilité à passer une commande
11. Montant des frais de port et d'emballage
12. Conformité entre photo et réalité des produits
13. Attrait, qualité et diversité des cadeaux
14. Possibilité d'essayer les produits
15. Prix des produits
16. Diversité et renouvellement des produits
17. Qualité et efficacité des produits
18. Attrait et caractère innovant des produits
19. Attrait et clarté des offres promotionnelles
20. Relation téléphonique pour renseignements
21. Relation téléphonique pour commande
22. Prise de contact en cas de problèmes



Boîte de dialogue de définition de l'analyse

Chaque thème est codé de 1 à 5 :

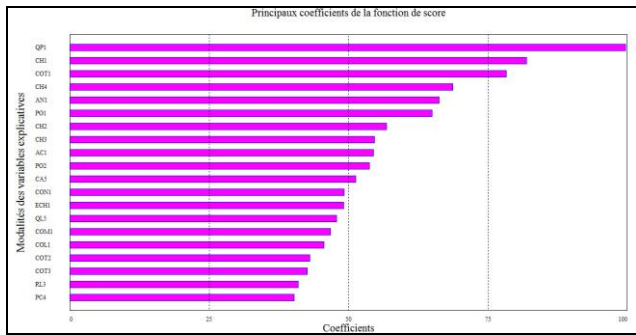
- 1 = Très satisfaite
- 2 = Satisfaite
- 3 = Indifférente
- 4 = Mécontente
- 5 = Très mécontente

L'âge des clientes est codé de la façon suivante :

- 1 = plus de 50 ans
- 2 = 30 à 50 ans
- 3 = moins de 30 ans

La satisfaction des clientes est codée de la façon suivante : 1 = Très satisfaite, 2 = Autres (satisfaite, indifférente, mécontente, très mécontente)

Après avoir mis en œuvre la procédure, divers graphiques sont disponibles. Le premier affiche les plus grands coefficients de la fonction de score. Il permet de visualiser les modalités des variables qui ont les plus forts impacts sur les scores des clientes.



Coefficients de la fonction de score

Le deuxième graphique affiche les deux courbes de répartition des scores des clientes « très satisfaites » et des « autres » clientes.

La zone rouge correspond aux clientes des deux groupes déclarées « autres » par la fonction de score, la zone orange correspond aux clientes déclarées « non classées » et la zone verte aux clientes déclarées « très satisfaites ».

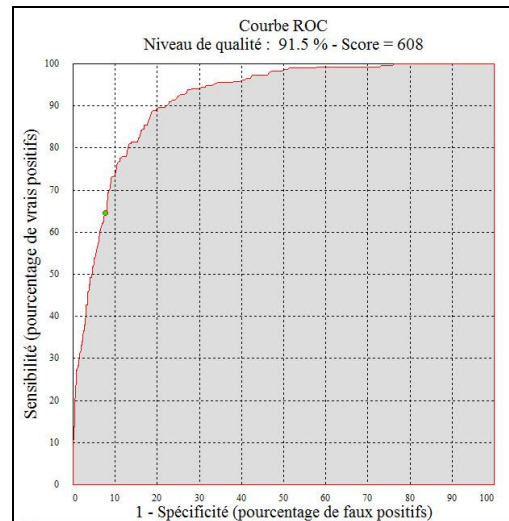
Ainsi les seuils de score sont définis pour des taux d'erreur de classement choisis par l'analyste.



Fonctions de répartition avec zone d'indécision

Dans certains cas, il est souhaité ne pas afficher de zone d'indécision. Cela est également proposé par cette procédure.

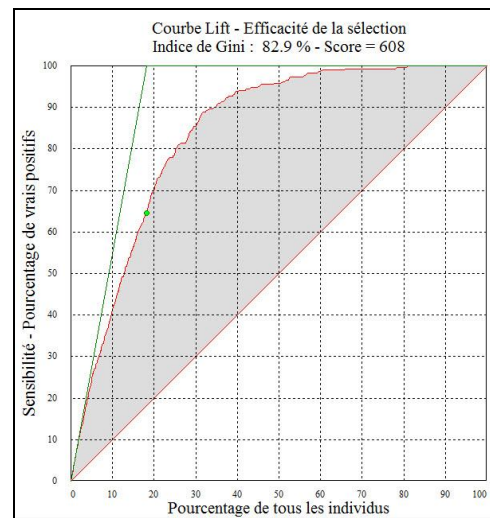
La courbe ROC (Receiver Operating Characteristic) affiche la relation entre la sensibilité (proportion de vrais « très satisfaites » chez les « très satisfaites » ou vrais positifs), et $1 -$ la spécificité (proportion de fausses « très satisfaites » chez les « autres » ou faux positifs). Plus cette courbe est proche de la partie supérieure du cadre du graphique, meilleure est la séparation entre les groupes réalisée par la fonction de score. Le niveau de qualité affiché représente l'aire sous la courbe.



Courbe ROC

La courbe de Lift affiche la relation entre la sensibilité et le pourcentage de tous les individus. La courbe Lift idéale est affichée en vert dans le graphique.

Cette courbe, est souvent utilisée dans le milieu de marketing. Elle permet notamment de répondre aux questions suivantes : avec un budget de campagne permettant de cibler un certain pourcentage de clientes, quelle sera la proportion d'acheteuses ? Une autre façon de poser le problème consiste à évaluer le pourcentage de clientes à cibler, donc les moyens à prévoir, afin d'atteindre une certaine proportion des acheteuses habituelles.



Courbe de Lift

L'indice de Gini représente la proportion grisée de l'aire entre la diagonale et la courbe Lift idéale.

Distributeur officiel : SIGMA PLUS
6, rue Collange - 92300 Levallois-Perret
Tél : 33(0) 1 47 30 37 30 – Fax : 33 (0)1 47 30 37 31
WWW.SIGMAPLUS.FR