Université de Carthage

Ecole Supérieure de la Statistique et Analyse de l’Information

Module : Econométrie

Classe : 2ème Année

Enseignant : Mokhtar KOUKI

Contact : mokhtar.kouki@essai.u-carthage.tn

Projet : « machine learning »

Année universitaire : 2021/2022

On considère un extrait de la base de données « German Credit Dataset » dont une description est fournie en annexe. Il s’agit de construire un modèle de prédiction de la probabilité de solvabilité du client. Il est demandé de suivre les étapes suivantes :

1. Réaliser une analyse univariée de toutes les variables de la base de données
2. Réaliser une analyse bivariée entre la variable cible (Statut) et toutes les variables explicatives
3. Construire un modèle prédictif d'une base d'apprentissage (train) en mettant en compétition plusieurs méthodes
4. Fournir des indicateurs de performances des prévisions réalisées sur la base test
5. Réaliser une application RShiny qui reprend les points 1-4)
6. Inclure dans l’application RShiny un formulaire qui permet de saisir les informations d’un client et qui a pour finalité d’estimer la probabilité que le client soit solvable.

Annexe

Description de la « German Credit Dataset »

**Solde du compte courant**

1 : ... < 0

2 : 0 <= ... < 200

1. : ... >= 200
2. : Pas de compte courant

**Durée du crédit en mois**

**Historique des crédits :**

1. : aucun crédit pris / tous les crédits remboursés
2. : tous les crédits de cette banque ont été remboursés
3. : crédits déjà remboursés jusqu'à maintenant
4. : retard dans le paiement dans le passé
5. : compte critique / autres crédits existants (pas dans cette banque)

**Motif**

1. : Voiture Nouvelle
2. : Voiture d’occasion
3. : fourniture/équipement
4. : Télévision
5. : Appareil électroménager
6. : Réparation
7. : Education
8. : Recyclage (reconversion)
9. : Projet
10. : Autres

**Montant du crédit**

**Compte d’épargne/Obligation**

1: ... < 100

2: 100 <= ... < 500

3: 500 <= ... < 1000

4: .. >= 1000

5: Pas de compte épargne

**Actuellement occupé depuis :**

* 1. : En chômage
  2. : ... < 1 ans
  3. : 1 <= ... < 4 ans
  4. : 4 <= ... < 7 ans

5 : .. >= 7 ans

**Statut personnel et sexe :**

1 : Homme : Divorcé

2 : Femme : Divorcée/Mariée

3 : Homme : Célibataire

4 : Homme : Marié/Veuf

5 : Femme : Célibataire

**Autres débiteurs / garants**

* 1. : Aucun
  2. : Codemandeur
  3. : Garant

**Résidence actuelle depuis**

**Age**

**Logement**

* 1. : Locataire
  2. : Propriétaire
  3. : Gratuitement

**Nombre de crédits en cours dans cette banque**

**Emploi :**

1. : sans emploi / non qualifié - non-résident
2. : non qualifié - résident
3. : Employé/Fonctionnaire
4. : Manager/Indépendant/Ouvrier Qualifié

**Nombre de personnes à charge**

**Statut:**

1: Solvable

2: Non solvable