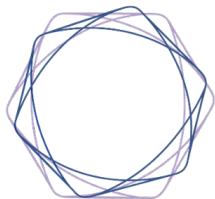


Outil d'interprétation de score

B. Thieurmel, Datastorm
7e Rencontres R - Rennes
5 juillet 2018



DataStorm

Motivations

- Utilisation de plus en plus fréquente de méthodes de **Machine Learning, d'IA, et de Deep Learning**^a dans le cadre de nos projets

^aForêts aléatoires, xgboost, réseaux de neurones, ...

Problème ?

- Scores "*boite noire*"
- ... mais la volonté de pouvoir les interpréter !

Solution

- Développement d'un outil d'aide à l'interprétation d'un score *via* une application **shiny**

Point de départ : un jeu de données avec la variable réponse, des variables explicatives (prises en compte ou non dans la modélisation initiale) ainsi que le score obtenu.

Questions :

- Quelle est la performance du score *au global* ?
- Quel choix faire pour la valeur du seuil ?
- Quelles sont les variables qui ont le plus contribué à l'élaboration du score ?
- Peut-on expliquer ou catégoriser les bonnes ou les mauvaises prédictions ?
- Quelle est la répartition du score et sa performance selon les valeurs d'une des variables explicatives ?

Chargement des données

The screenshot shows the 'Interpretation Toolbox' interface. On the left, there is a sidebar with sections: 'Report data' (with 'RDS or csv' selected), 'Subset?' (with 'Bon_nouveaux_payeur' selected), 'Choose subset variable' (with 'Bon_nouveaux_payeur' selected), 'Choose modalities' (with 'Bon_nouveaux' selected), 'Score', 'Documentation', and 'About'. The main area displays a dataset table with the following columns: 'ban', 'compte courant', 'date', 'historique credit', 'objet credit', 'montant credit', 'compte epargne', 'duree dernier emploi', 'pourcentage variation salaire', 'autre detendeur garant', 'duree dernière résidence', 'propriétés', 'age', 'autres comptes provisions/ma', and 'logement'. The table contains 5 rows of data. At the bottom, it says 'Showing 1 to 6 of 1,993 entries' and 'Summary'.

- .RDS ou .csv
- Sous-ensemble de lignes ?
- Table + *summary*

Données de démonstration

- Statlog (German Credit Data) Data Set ^a
- 1 000 individus, 7 variables quantitatives, 13 qualitatives
- Modélisation du risque de défaut avec des *forêts aléatoires* sur l'ensemble des variables

^a[https://archive.ics.uci.edu/ml/datasets/statlog+\(german+credit+data\)](https://archive.ics.uci.edu/ml/datasets/statlog+(german+credit+data))

Performance du score

- Définition du score (*variable réponse, score, réponse positive, poids*)



- Répartition des réponses observées en fonction des tranches du score (déciles, quartiles, histogramme)
- Densités du score sur les sous-populations

Performance du score

Confusion matrix



Observed	Predicted		Class Error Rate
	Positive	Negative	
Positive	647	53	0.0757
Negative	184	116	0.6133

0.8004

AUC performance



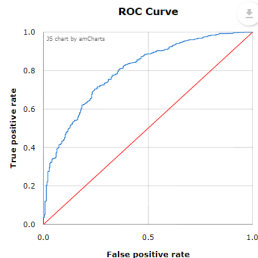
0.7786

Precision



ROC

Lift Precision Recall Precision/Recall Accuracy



Aide au choix du seuil

- Matrice de confusion **réactive** en fonction du seuil
- *Précision*, *AUC*, taux de bien classés
- Courbes *ROC*, *Lift*, *Precision/Recall*, ...

Interprétation du score

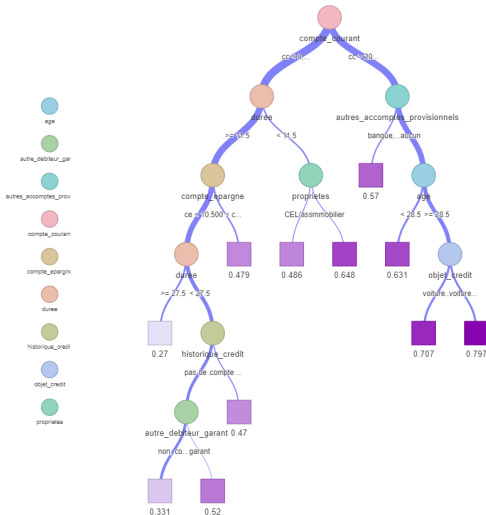
Idée (simple et visuelle...)

- Arbre de régression / classification expliquant au choix :
 - la vraie réponse observée
 - le score modélisé
 - la qualité globale ou détaillée (*TP, FP, TN, FN*)
- Analyse et visualisation interactive de l'arbre. Détection et catégorisation de *niches*
- Performance de l'arbre (*Train/Test* possible)
- Importance des variables et croisement avec le score et les réponses

The screenshot shows a web-based interface for analyzing a decision tree. It includes several sections:

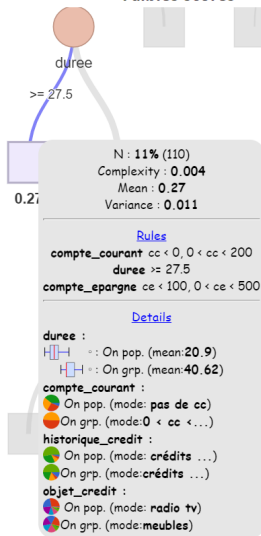
- Parameters:** A header for the configuration options.
- Choose explanatory variables:** A list of variables including `compte_courant`, `duree`, `historique_credit`, `objet_credit`, `montant_credit`, `compte_epargne`, `duree_dernier_emploi`, `pourcentage_versement_salaires`, `auvre_debiteur_garant`, `duree_derniere_residence`, `propreties`, `age`, `autres_comptes_provisionnels`, `logement`, `nb_credits`, `emploi`, `nb_pers_maintenance`, `telephone`, `travailleur_etranger`, and `sex`.
- Add graphics in tooltip:** A list of options including `duree`, `compte_courant`, and `historique_credit`.
- Focus on:** A dropdown menu with options: `Real response`, `Real response`, `Score`, `Quality predictions (global)`, and `Quality predictions (detailed)`.
- Split in train/test?** A checkbox that is checked.
- Seed:** A text input field containing the value 123.
- % Train:** A text input field containing the value 80.
- % Test:** A text input field containing the value 20.
- tree complexity:** A slider ranging from 0 to 0.11, with a blue bar indicating the current setting.
- tree minsplit:** A slider ranging from 0 to 200, with a blue bar indicating the current setting.
- Interpolate:** A button located at the bottom right of the interface.

Focus sur le score : arbre et performance



Focus sur le score : première analyse descriptive

Faibles scores

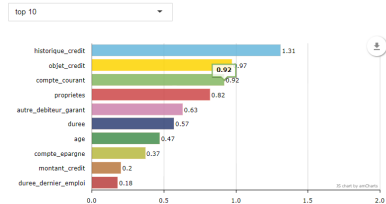


Forts scores



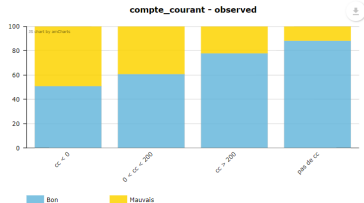
Focus sur le score : importance des variables

Variables importance



Modality	N	Accuracy	Real Pos. (%)	Of which pred. Pos. (%)	Real Neg. (%)	Of which pred. Neg. (%)
cc < 0	274	55.47%	50.73%	13.67%	49.27%	86.52%
0 < cc < 200	269	46.84%	60.97%	15.24%	39.03%	96.19%
cc > 200	63	41.27%	77.78%	36.73%	22.22%	57.14%
pas de cc	394	76.14%	88.32%	78.74%	11.68%	52.17%

Count

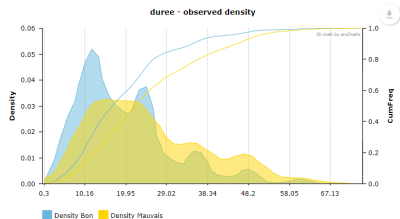


- un clique sur la variable pour avoir un focus sur son impact sur le score
- ici, beaucoup de mauvais payeurs pour $cc < 0$, qui sont bien prédits par le modèle
- retour sur les données brutes pour observer cette relation

Focus sur le score : importance des variables

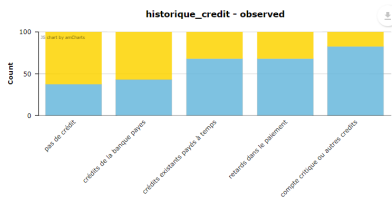
duree

Modality	N	Accuracy	Real Pos. (%)	Of which pred. Pos. (%)	Real Neg. (%)	Of which pred. Neg. (%)
[0, 5)	6	100.00%	100.00%	100.00%	0.00%	0.00%
[5, 10)	137	81.02%	82.48%	86.73%	17.52%	54.17%
[10, 15)	224	74.11%	76.34%	76.02%	23.66%	67.92%
[15, 20)	179	64.80%	69.27%	63.71%	30.73%	67.27%
[20, 25)	224	73.66%	70.54%	70.25%	29.46%	81.82%
[25, 30)	17	70.59%	64.71%	63.64%	35.29%	83.33%
[30, 35)	43	60.47%	67.44%	55.17%	32.56%	71.43%
[35, 40)	88	63.64%	56.82%	46.00%	43.18%	96.84%
[40, 45)	12	50.00%	66.67%	37.50%	33.33%	75.00%
[45, 50)	54	68.52%	40.74%	36.36%	59.26%	90.63%

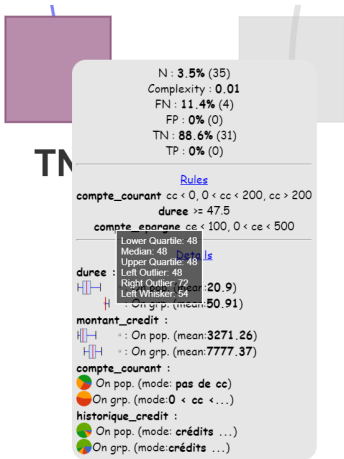


historique_credit

Modality	N	Accuracy	Real Pos. (%)	Of which pred. Pos. (%)	Real Neg. (%)	Of which pred. Neg. (%)
pas de crédit	40	70.00%	37.50%	33.33%	62.50%	92.00%
crédits de la banque payés	49	69.39%	42.86%	42.86%	57.14%	89.29%
crédits existants payés à temps	530	68.68%	68.11%	64.27%	31.89%	78.11%
retards dans le paiement	88	59.09%	68.18%	55.00%	31.82%	67.86%
compte critique ou autres crédits	293	90.78%	82.94%	85.19%	17.06%	58.00%



Focus sur la qualité du modèle : arbre et performance



Pour conclure

Tenté par un essai ? Application démo en ligne

https://datastorm-open.shinyapps.io/tools_interpretation/

Next steps

- Prise en compte des modèles de régression
- Méthode(s) plus robuste(s) que de simples arbres CART pour l'importance des variables

Nos packages R

- Développés par les équipes de Datastorm sur leur temps de R&D
- Disponibles sur le CRAN
- **visNetwork** : visualisation interactive de réseaux et d'arbres CART
- **rAmCharts** (graphiques js avec *amCharts*), **suncalc**, **ROI.plugin.clp**

<https://github.com/datastorm-open>