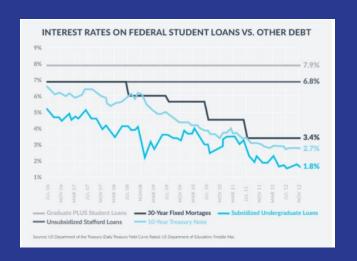
Veritas Softwares

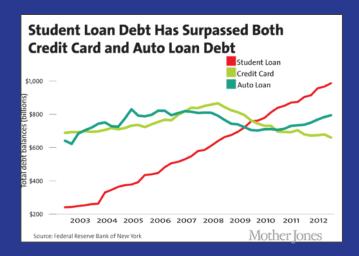
Prédicteur de salaire : SalaPredi™



8 décembre 2017 FERRÉ Nicolas, MOLLARD Romaric, RUCHOT Guillaume

Contexte

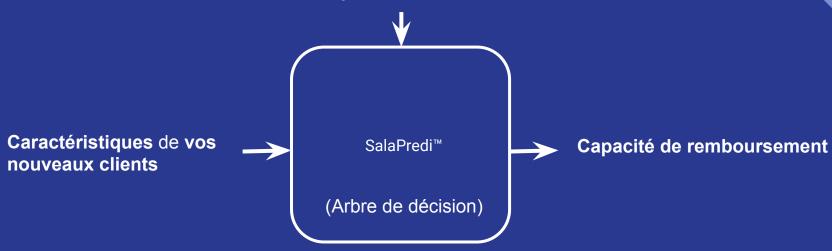




> Assurer le remboursement des prêts de manière prévisionnelle <



Caractéristiques de vos clients courants





Caractéristiques de vos clients courants

> Hypothèses:

- Les données incomplètes ne sont pas déterminantes
- Il nous faut des données avec le salaire pour l'apprentissage

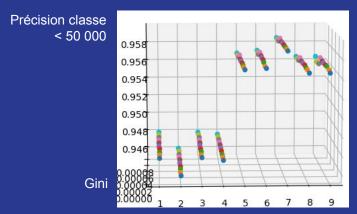
> Choix:

- Suppression des attributs trop peu connus
- Suppression des données restantes incomplètes
- Formatage automatique des données



SalaPredi™

Optimisation des paramètres



Profondeur de l'arbre

Paramètres calculés automatiquement pour ces données :

Gini : 0.00 006 profondeur : 7





Réel \ Prédiction	< 50 000	> 50 000
< 50 000	93%	1%
> 50 000	4%	2%

Paramètres calculés automatiquement pour ces données :

Gini: 0.00 006 profondeur: 7





Précision par classe

Précision de l'affirmation > 50 000 : **70%** Précision de l'affirmation < 50 000 : **96%**

Performances

20 000 prédictions par secondes Apprentissage (300 000 entrées) optimisé en 30 minutes

Paramètres calculés automatiquement pour ces données :

Gini : 0.00 006 profondeur : 7



Limitations

- Format de données
- Les données doivent être complètes
- Disparité entre les deux classes

Améliorations

- + Adaptation aux bases de données
- + Test d'autres modèles
- + Interface utilisateur et simplifications



Conclusion

- Automatisation du traitement de l'entrée avec filtrage des données
- Paramétrage automatique
- Paramétrage manuel possible (prétraitement, apprentissage)
- Résultats précis dans une certaine mesure (70% et 96%)

