

## Prix LITRA 2023 : Résumé du travail primé

Auteur :	Zeliya Schär
Titre :	Vorhersage des Upselling-Potenzials von Kundinnen und Kunden des öffentlichen Verkehrs mittels Supervised Machine Learning
Date :	02 juin 2023
Établissement :	Haute École de Lucerne – Économie
Filière :	Master of Science in Applied Information and Data Science
Niveau :	Travail de master

### Résumé

Le BLS veut augmenter la vente de billets de première classe et de surclassements, et donc son chiffre d'affaires. Dans ce travail, l'auteure cherche à savoir si l'apprentissage automatique supervisé peut aider à affiner de manière ciblée la sélection des clients lors des campagnes de marketing. Certains modèles sont en effet capables de prédire des clients potentiels à partir de données de transaction anonymisées. D'où leur intérêt pour développer des mesures de marketing spécifiques aux groupes cibles et éviter les pertes de diffusion.

### Situation initiale

Le BLS est raccordé au système tarifaire et de distribution de la plateforme nationale NOVA depuis 2020. Le système enregistre donc toutes les transactions effectuées par ses clients. L'entreprise de transports souhaite analyser les données de transactions anonymisées existantes pour ensuite implémenter les éventuels résultats dans son marketing opérationnel afin d'augmenter son chiffre d'affaires par la vente accrue de billets de première classe et de surclassement. Afin de mieux exploiter le potentiel de montée en gamme, il est important de connaître les groupes cibles pour qui de telles offres sont attrayantes. L'utilisation d'un modèle d'apprentissage automatique supervisé peut augmenter le succès du marketing en permettant de présélectionner les clients de manière ciblée et de s'adresser à eux personnellement.

### Sujet de recherche

Avec quelle précision l'apprentissage automatique supervisé permet-il de prédire le potentiel de montée en gamme chez les clients du BLS résidant en Suisse sur la base des données de transactions anonymisées des années 2020 à 2022 ? Quels sont les facteurs exerçant la plus forte influence sur ce potentiel ?

### Démarche

La démarche suit le modèle de processus d'exploration des données « Cross Industry Standard Process for Data Mining » (CRISP-DM) établi pour le « data mining ». Zeliya Schär décrit tout d'abord le business case à étudier. Après avoir préparé les données mises à disposition par le BLS à partir de la base de données des transactions anonymisées fournie par NOVA, elle procède à leur analyse. Le jeu de données obtenu est ensuite complété et nettoyé. Le cœur du travail consiste à appliquer l'apprentissage automatique, c'est-à-dire à reconnaître, à l'aide de l'informatique, des modèles pour différentes

---

caractéristiques. Pour ce faire, plusieurs algorithmes ou modèles sont alimentés et « entraînés » avec des jeux de données de test. Les résultats sont comparés à des résultats aléatoires au moyen de différents tests et la précision des prévisions est évaluée. L'auteur applique cette procédure à quatre algorithmes, qu'elle alimente chacun avec deux ou trois jeux de données traités différemment. Pour finir, elle évalue les résultats et en déduit des conclusions sur le potentiel de montée en gamme.

### Résultats

Parmi les quatre algorithmes, « Random Forest » affiche les prévisions les plus précises avec un taux de 90 %. Dans le jeu de données test, la méthode reconnaît plus des trois quarts des clients à être déjà montés en gamme. En outre, elle est capable d'identifier le groupe de personnes qui pourrait être réceptif à cette perspective. Dans les quatre modèles, l'âge, le revenu ainsi que le nombre et le type de titres de transport achetés s'avèrent être des prédicteurs essentiels. Ainsi, la probabilité d'une montée en gamme augmente dans les cas suivants : Lorsqu'une personne a acheté plus d'un ticket, lors de l'achat de tickets individuels, lorsqu'une personne est âgée de plus de 28 ans et lorsqu'une personne possède un abonnement. De même, les personnes disposant d'un revenu plus élevé ont davantage tendance à effectuer des ventes incitatives.

*Auteur du résumé à la demande de la LITRA : Rémy Chrétien, Federas Beratung AG  
Juillet/août/septembre 2023*