

## Prix LITRA 2023: Zusammenfassung prämierte Arbeit

<b>Autorin:</b>	Zeliya Schär
<b>Titel:</b>	Vorhersage des Upselling-Potenzials von Kundinnen und Kunden des öffentlichen Verkehrs mittels Supervised Machine Learning
<b>Datum:</b>	02.06.2023
<b>Institution:</b>	Hochschule Luzern (HSLU) – Wirtschaft
<b>Studiengang:</b>	Master of Science in Applied Information and Data Science
<b>Level:</b>	Masterarbeit

### Zusammenfassung

Die BLS AG will den Verkauf von Erstklassbilleten und Klassenwechseln und damit ihren Umsatz steigern. Diese Arbeit beantwortet die Frage, wie Supervised Machine Learning helfen kann, die Kundenselektion bei Marketingkampagnen zielgerichtet zu verfeinern. Tatsächlich sind einige Modelle in der Lage, aus anonymisierten Transaktionsdaten potenzielle Kundinnen und Kunden vorherzusagen. Damit sind sie wertvoll, um zielgruppenspezifische Marketingmassnahmen zu entwickeln und Streuverluste zu vermeiden.

### Ausgangslage

Seit 2020 ist die BLS an das schweizweite Preis- und Vertriebssystem, NOVA, angeschlossen. Dieses System erfasst seither alle Transaktionen der BLS-Kundschaft. Die BLS AG möchte die vorhandenen anonymisierten Transaktionsdaten analysieren und mögliche Ergebnisse anschliessend im operativen Marketing implementieren, um ihren Umsatz mit dem vermehrten Verkauf von Erstklassbilleten und Klassenwechseln zu erhöhen. Um das Potenzial dieses Upsellings besser zu nutzen, ist es wichtig zu wissen, für welche Zielgruppen derartige Angebote attraktiv sind. Der Einsatz eines Supervised Machine Learning-Modells kann den Marketingerfolg erhöhen, indem er ermöglicht, Kundinnen und Kunden gezielt zu selektionieren und persönlich anzusprechen.

### Forschungsfrage

Mit welcher Genauigkeit kann das Upselling-Potenzial von in der Schweiz wohnhaften Kundinnen und Kunden der BLS AG aufgrund von anonymisierten Transaktionsdaten der Jahre 2020 bis 2022 mit Supervised Machine Learning vorhergesagt werden und durch welche Faktoren wird dieses Potenzial am stärksten beeinflusst?

### Vorgehen

Das Vorgehen folgt dem für Data Mining etablierten Prozessmodell «Cross Industry Standard Process for Data Mining» (CRISP-DM). Zunächst beschreibt Zeliya Schär den zu untersuchenden Business Case. Sie bereitet dann die von der BLS zur Verfügung gestellten anonymisierten Daten aus der Transaktionsdatenbank NOVA auf, führt eine Datenanalyse durch und ergänzt und bereinigt den Datensatz. Der Hauptteil der Arbeit besteht darin, das Machine Learning anzuwenden, das heisst rechnergestützt Muster für verschiedene Merkmale zu erkennen. Zu diesem Zweck werden mehrere Algorithmen resp. Modelle mit Testdatensätzen gefüttert und damit «trainiert». Die Ergebnisse werden

---

mittels verschiedener Tests mit Zufallsergebnissen verglichen und die Prognosegenauigkeit beurteilt. Dieses Verfahren wendet die Autorin auf vier Algorithmen mit jeweils zwei bis drei unterschiedlich bearbeiteten Datensätzen an. Zum Schluss evaluiert sie die Ergebnisse und leitet daraus Folgerungen für das Upselling-Potenzial ab.

### Ergebnisse

Unter den vier Algorithmen zeigt «Random Forest» mit 90% die beste Prognosegenauigkeit. Das Modell erkennt im Testdatensatz mehr als drei Viertel der Kundinnen und Kunden, die bereits ein Upselling gemacht haben. Ausserdem ist es in der Lage, die Personengruppe zu identifizieren, die für ein Upselling empfänglich sein könnte. In allen Modellen erweisen sich das Alter, das Einkommen sowie die Anzahl und Art gekaufter Fahrausweise als wesentliche Prädiktoren. So steigt die Wahrscheinlichkeit für ein Upselling in folgenden Fällen: Wenn eine Person mehr als ein Ticket gekauft hat, der Kauf von Einzeltickets, wenn eine Person über 28-jährig ist und wenn eine Person ein Abonnement besitzt. Ebenso tendieren Personen mit einem höheren Einkommen eher zu einem Upselling.

*Autor der Zusammenfassung im Auftrag der LITRA: Rémy Chrétien, Federas Beratung AG  
Juli/August/September 2023*