

LU02.A02 - Fallstudie Machine Learning

Lernziele

1. Ablauf von Machine Learning anwenden können.

Rahmenbedingungen

- Sozialform: Partnerarbeit
- Zeitbudget: 40 Minuten
- Erwartetes Ergebnis: Python Script, dass auf Input-Daten Kaufentscheidungen des Kunden *vorhersagen* kann.

Ausgangslage

Ein Online-Shop möchte vorhersagen, ob ein Kunde ein Produkt kaufen wird oder nicht. Gegeben seien die nachfolgenden Attribute:

- age (Alter)
- past_purchases (Anzahl bisheriger Einkäufe)
- minutes_on_page (Verweildauer)

* buy (Zielvariable: 0/1)

Auftrag A - Datensatz erstellen - 10'

1. Erstelle Sie in Python eine Datei shop_data.csv mit mindestens 30 Zeilen.
2. Verwenden Sie die nachfolgenden plausible Wertebereiche:

```
* age: 16–70
* past_purchases: 0–30
* minutes_on_page: 0.1–20
* buy: 0 oder 1
```

3. Bauen Sie eine einfache "Logik" ein, damit das Problem nicht reines Raten ist (z. B. höherer Kauf bei mehr Verweildauer + mehr bisherigen Einkäufen).

Auftrag B - ML-Pipeline bauen

Schreiben Sie ein Skript ml_basics_shop.py, das:

- die CSV einliest (pandas)
- Features (X) und Zielvariable (y) trennt
- Daten in Train/Test splittet (z. B. 80/20, random_state=42)
- eine Pipeline erstellt mit:
 - Standardisierung (StandardScaler)
 - Modell: Logistische Regression
- das Modell trainiert
- Vorhersagen fuer Testdaten berechnet

Teil C - Bewertung

Geben Sie die nachfolgende Kennzahlen aus:

- * Accuracy
- * Confusion Matrix
- * Classification Report (Precision/Recall/F1)

Teil D - Zusatz

Zusatz: Erkläre in 2-3 Sätzen, warum Accuracy allein manchmal irreführend ist.

Lösung

[Lösung](#)



Volkan Demir

From:

<https://wiki.bzz.ch/> - **BZZ - Modulwiki**

Permanent link:

<https://wiki.bzz.ch/de/modul/m245/learningunits/lu02/aufgaben/02?rev=1767615176>

Last update: **2026/01/05 13:12**

