

Verlustprävention an Selbstbedienungskassen

Real Data Mining GmbH

April 2025

Inhaltsverzeichnis

- 1 Ausgangslage & Zielsetzung
- 2 Datenlage
- 3 Bewertungslogik
- 4 Projektphasen
- 5 Risiken & Ausblick

- Stationäre SBK ohne Scanner-Abgleich
- Volle Verantwortung beim Kunden
- Bisher: Stichprobenkontrollen, keine Systematik

Was wir erreichen wollen

Entwicklung eines Algorithmus zur Erkennung verdächtiger Transaktionen zur gezielten Nachkontrolle

- CSV: Stammdaten, Parquet: Transaktionen
- Zeitraum: 2022–2024
- Gelabelt: label, damage

- Nur 3,14% Betrugsanteil unter kontrollierten Transaktionen
- Teilweise fehlende Werte (camera, feedback)
- Strukturfehler: `n_lines` inkonsistent

Bewertung nach Kundenvorgabe

Fall	Wert [€]
Keine Kontrolle, korrekt	0,00
Keine Kontrolle, Betrug	– Warenwert
Kontrolle, korrekt	–10,00
Kontrolle, Betrug	+5,00

Bewertungserweiterung (Vorschlag)

- Parametrisierbar statt fest
- Kontextabhängig: Artikel, Uhrzeit, Filiale
- Reputations- & Kundenfaktor berücksichtigbar

Meilensteinplan (DASC)

Nr	Inhalt	Phase	Start	Ende
1	Zielklärung, Projektauftrag	Define	07.04.2025	29.04.2025
2	Datenanalyse, REST-Vorgabe	Acquire	07.04.2025	18.05.2025
3	Feature Engineering, Modell	Structure	19.05.2025	15.06.2025
4	Bewertung, REST-Prototyp	Communicate	16.06.2025	29.06.2025

- Unausgewogene Klassen (Imbalance)
- Fehlende Dokumentation (kein Data Dictionary)
- Generalisierbarkeit auf andere Filialen unklar

- Modell kann Kontrollaufwand ökonomisch steuern
- Kombination von Klassifikation + Regressionsanalyse zielführend
- REST-Schnittstelle ermöglicht Integration



Vielen Dank! Fragen?